

Advances in Sociophonetics

Edited by
Chiara Celata
Silvia Calamai

John Benjamins Publishing Company

Studies in Language Variation

15

Advances in Sociophonetics

Studies in Language Variation

The series aims to include empirical studies of linguistic variation as well as its description, explanation and interpretation in structural, social and cognitive terms. The series will cover any relevant subdiscipline: sociolinguistics, contact linguistics, dialectology, historical linguistics, anthropology/anthropological linguistics. The emphasis will be on linguistic aspects and on the interaction between linguistic and extralinguistic aspects — not on extralinguistic aspects (including language ideology, policy etc.) as such.

For an overview of all books published in this series, please see <http://benjamins.com/catalog/slv>

Editors

Peter Auer
Universität Freiburg

Frans Hinskens
Meertens Instituut &
Vrije Universiteit,
Amsterdam

Paul Kerswill
University of York

Editorial Board

Jannis K. Androutsopoulos
University of Hamburg

Peter Gilles
University of Luxembourg

K. K. Luke
The University of Hong Kong

Arto Anttila
Stanford University

Barbara Horvath
University of Sydney

Rajend Mesthrie
University of Cape Town

Gaetano Berruto
Università di Torino

Brian Joseph
The Ohio State University

Pieter Muysken
Radboud University Nijmegen

Paul Boersma
University of Amsterdam

Johannes Kabatek
Eberhard Karls Universität
Tübingen

Marc van Oostendorp
Meertens Institute & Leiden
University

Jenny Cheshire
University of London

Juhani Klemola
University of Tampere

Sali Tagliamonte
University of Toronto

Gerard Docherty
Newcastle University

Miklós Kontra
University of Szeged

Johan Taeldeman
University of Gent

Penny Eckert
Stanford University

Bernard Laks
CNRS-Université Paris X
Nanterre

Øystein Vangsnes
University of Tromsø

William Foley
University of Sydney

Maria-Rosa Lloret
Universitat de Barcelona

Juan Villena Ponsoda
Universidad de Málaga

Volume 15

Advances in Sociophonetics

Edited by Chiara Celata and Silvia Calamai

Advances in Sociophonetics

Edited by

Chiara Celata

Scuola Normale Superiore, Pisa

Silvia Calamai

Università degli Studi di Siena

John Benjamins Publishing Company

Amsterdam / Philadelphia



The paper used in this publication meets the minimum requirements of the American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

Advances in Sociophonetics / Edited by Chiara Celata and Silvia Calamai.

p. cm. (Studies in Language Variation, ISSN 1872-9592 ; v. 15)

Includes bibliographical references and index.

1. Grammar, Comparative and general--Phonology. 2. Language and languages--Variation. 3. Dialectology--Research. 4. Sociolinguistics. I. Celata, Chiara, 1979--editor of compilation. II. Calamai, Silvia, editor of compilation.

P217.3.A28

2014

417--dc23

2013050139

ISBN 978 90 272 3495 7 (Hb ; alk. paper)

ISBN 978 90 272 7050 4 (Eb)

© 2014 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands

John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

Table of contents

Introduction: Sociophonetic perspectives on language variation <i>Chiara Celata and Silvia Calamai</i>	1
Part I. Variation and sociolinguistics	
CHAPTER 1	
The sociophonetic orientation of the language learner <i>William Labov</i>	17
CHAPTER 2	
French liaison and the lexical repository <i>Bernard Laks, Basilio Calderone and Chiara Celata</i>	31
Part II. Sources and functions of sociophonetic variation	
CHAPTER 3	
Derhoticisation in Scottish English: A sociophonetic journey <i>Jane Stuart-Smith, Eleanor Lawson and James M. Scobbie</i>	59
CHAPTER 4	
Where and what is (t,d)? A case study in taking a step back in order to advance sociophonetics <i>Rosalind A. M. Temple</i>	97
CHAPTER 5	
New parameters for the sociophonetic indexes: Evidence from the Tuscan varieties of Italian <i>Giovanna Marotta</i>	137
CHAPTER 6	
Sound archives and linguistic variation: The case of the Phlegraean diphthongs <i>Rosanna Sornicola and Silvia Calamai</i>	169

Part III. What is (and what is not) a sociophonetic change

CHAPTER 7

Ejectives in English and German – Linguistic, sociophonetic, interactional, epiphenomenal?	189
<i>Adrian Simpson</i>	

Author index	205
---------------------	------------

Subject index	209
----------------------	------------

Sound and video files for Chapters 3 and 4
can be found at
<http://dx.doi.org/10.1075/silv.15.media>

Introduction

Sociophonetic perspectives on language variation

Chiara Celata* and Silvia Calamai**

* Scuola Normale Superiore di Pisa

** Università degli Studi di Siena

1. Why this volume?

This volume collects seven papers in contemporary sociophonetic research. It addresses hot themes in sociophonetics and proposes a fresh look at old problems still open to debate. A variety of approaches is proposed without neglecting the need for a coherent discussion of the nature of variation in speech and how speakers develop a cognitive representation of it. These characteristics distinguish the present volume from the panorama of comparable sociophonetic literature, which mainly consists of textbooks, readers, and journal special issues (as well as individual journal articles, conference proceedings, and informal reports).

According to Jannedy & Hay (2006), contemporary sociophonetics and sociophonology differ from early variationist sociolinguistics for their focus on the cognitive representation of phonetic variation in the mind of the individual. Stated differently, the fundamental purpose of sociophonetic studies should be that of analyzing how the concrete communicative experiences are categorized by the speakers and, most importantly, of establishing the function of such complex nucleus of information in the structuring of linguistic systems. The fusion of sociolinguistics and phonetics occurs therefore within a cognitivist perspective in which the probabilistic nature of the language and the interest for the processes of language use and comprehension play a special role. The diffused reference to usage-based models of language perception, production and representation and to Exemplar Theory (Goldinger 1998; Bybee 2000; Pierrehumbert 2001) possibly is the most salient cue of the discontinuity from early accounts of sociolinguistic variation (see, for a similar reasoning, the paper by Laks et al. in this volume). Exemplar Theory is considered a resource for understanding and modelling the

dynamics of acquisition of socio-indexical variation (Foulkes 2010). Starting from the observation that socio-indexical variation, alongside physically- and mechanically-based phonetic variation, is not dysfunctional to, but rather favors language acquisition (a position which is not entirely new; see e.g. *Natural Phonology* and Donegan & Stampe 1979), sociophoneticians recognize that the exemplarist perspective is able to account for the fact that socio-indexical variation turns out to even be a prerequisite for the development of abstract categories from superficial exemplar stores.

The idea that the vast amount of speech variation experienced daily by children acquiring their native language and adult speakers and hearers is a challenge for any account of phonological processing can be seen as the principal unifying feature of the book. Most contributions agree on the importance of developing complex methods and procedures for inspection and quantification of large corpora of speech data; some of them address the problem of how to implement a fine-grained instrumental investigation of those subtle details of phonetic variation that are only unsystematically attested in large corpora of unplanned speech; others specifically point to the importance of teasing out the sociophonetic dimension of variation from other dimensions of variation in speech, which are not sociophonetic.

Sociophonetics emerges therefore as the privileged domain for the investigation of language variation and change. As a matter of fact, it is the combination of theoretical reflections and sophisticated techniques of analysis, both phonetic and statistical, which have not reached the same level of elaboration in other domains of language investigation that allows the researcher to disentangle step by step the role of individual factors (sociolinguistic, physiological, communicative-interactive etc.) in the multidimensional space of speech variation.

At the same time, contemporary sociophonetics acknowledges its significant debt to historical dialectology and linguistic geography. Some of the chapters contained in this book offer critical insights into the legacy of traditional variationist linguistics, linguistic geography, and dialectology for current accounts of sociophonetic variation. There is an apparent historical paradox, clearly emerging in some of the papers, in the fact that those dialectal domains that were the first targets of modern scientific dialectology (first of all, the Italo-Romance domain) are still almost completely unexplored with respect to the social components of the observed sound changes. This book contains some concrete efforts toward a possible renovation of such eminent dialectological tradition through a more systematic examination of the sociolinguistic dimensions of sound change.

The book collects a brilliant array of contributors. Some of them may be considered among the founders of modern sociophonetics. Their field of expertise spans from experimental phonetics to dialectology, from phonology to

sociolinguistics. The case studies proposed (covering both Germanic – English and German – and Romance languages – French and varieties of the Italo-Romance domain) will appeal to the international audience. The originality of many of the issues treated and the methodologically sound approaches by which some theoretical nodes of the discipline are addressed make this volume a valuable resource for scholars interested in speech variation, social uses of language, and phonological representations.

2. Setting the stage: Variationism and sociolinguistics

In reading William Labov's paper, entitled "The Sociophonetic orientation of the language learner", one immediately realizes that the sociolinguistic version of functionalism has departed drastically from the internalist view of the language. It is stated in the article that "the individual does not exist as a unit of linguistic analysis" and that the individual patterns of variation have to be addressed not *per se*, but "to the extent to which they respond to wider community patterns". This Copernican revolution in language studies is not new, as the author points out; its roots are to be found in the critiques of the autonomy of idiolects of Weinreich, Labov, and Herzog (1968). However, the paper reframes the question by analyzing the strategies of phonological learning which allow the speaker/hearer to cope with the idiosyncratic constructs attested in the speech input.

The paper provides evidence that children may or may not adopt the features of parental language, depending on how these features match the features of the speech community. Children may reject the patterns of parental language and conform to the patterns of the surrounding community instead, especially in richly stratified societies whose members belong to different social and dialectal groups. Linguists are aware of this cross-generational effect because – as the author says – they often experience these mismatches in the speech of their own children.

The paper by Labov is of interest for the study of non-pathological attrition under sociolinguistic pressure (Köpke 2004) as well as issues/questions of language contact in general. Language attrition studies are most often concerned with adult learners (typically, post-pubescent migrants) and the focus of the analysis is generally on the effects of "transitional" (or intra-generational) bilingualism generated by an L1-to-L2 gradual shift. On the contrary, the paper by Labov is strictly concerned with the speech of children as opposed to adult parental speech, thus adopting an intergenerational perspective. Moreover, the study deals with internally varied linguistic communities, where no specific variety seems to play the role of the "dominant" language. Labov's study suggests that the range of variation that a speaker may experience not only depends on input variability, but also on a smoothing

function operated by the linguistic community with respect to that variability. As the acquisitional evidence reviewed in the paper demonstrates, the language learner has a “compulsion [...] to turn outward” (p. 18) in the direction of the community patterns and successive contacts experienced in life after the formation of the initial competence may erode parents’ influence in a gradual but continuous way.

From this analysis the author concludes that the individual patterns are not as informative as is normally suggested by proponents of different types of “linguistic individualism”; it is rather the case that the larger the picture, the more informative the data therein.

The paper by Labov constitutes an excellent introduction to the rest of the papers collected in the volume. The emphasis on the accommodation mechanisms that individual learners carry out in the construction of individual grammars is but one of the many possible responses to the question of how the vast amount of speech variation with indexical meaning is cognitively represented – an issue that is directly or indirectly addressed by all of the papers in the volume.

In a similar vein, and taking its inspiration from usage-based models of grammar and exemplar theory (Goldinger 1998; Bybee 2001), the paper by Bernard Laks, Basilio Calderone and Chiara Celata “French Liaison and the lexical repository” starts from the well-known lexicalist hypothesis that liaison is more frequently realized in those word groups that have strong internal cohesion and high frequency of co-occurrence (Bybee 2005), and shows that this hypothesis, whose substantial correctness must be confirmed, should be refined in some respects if we take the procedures of corpus analysis seriously and investigate a very large amount of actual productions realized by different groups of French speakers.

The materials of the analysis are the 16,805 sites of realized liaisons coded in the PFC corpus (Durand et al. 2002). PFC is the largest database of spoken French currently available, and has been collected over many years according to the “Labovian” paradigm of sociolinguistic enquiries. Based on such large repository of actual uses, the paper shows that the liaison distribution is similar to a power-law distribution in which a few word junctures are ranked high for productivity and account for approximately one-half of the total observations, while a very long list of less productive or unproductive junctures accounts for the remaining half of the realizations. The authors argue that this statistical distribution goes beyond some traditional views of linguistic storage, according to which part of the liaison process must be inscribed as a nucleus of stored “constructions” in the mental lexicon while low frequency constructions tend to be lost (Bybee 2005). Quite on the contrary, liaison shows that storage is limited to a relatively short (but cognitively heavy) list of occurrences, while “a productive process of generalization” (p. 39) must account for the long tail of dispersed, low-frequency realizations.

In this approach, the authors appear to owe much to the view according to which the balance between storage and computation in language processing cannot be defined once and for all. The cost of storage is not necessarily more than the cost of processing (e.g., Baayen et al. 2002); in investigating corpora of actual uses, storage may be found to cover much of the labour necessary for processing specific phenomena, provided that a fine-grained analysis of frequency distributions is realized.

The proposed generalizations are found to hold true also for subgroups of data as defined by different types of liaison consonants (/n/, /z/, /t/) and the speakers' age and educational level. Age was chosen because it is known to be a relevant sociolinguistic factor in liaison variation (Durand et al. 2011); on the contrary, educational level was chosen because it has always been disregarded in previous analyses of French liaison. The results show that variation in liaison production as a function of educational level is present in the "tail" of the distribution, that is, in that variegated sample of low- and very-low-frequency items which clearly turns out to be "the most likely repository of lexical environments differentially selected by different groups of speakers" (p. 47). In the authors' opinions, such a result confirms the importance of adopting a corpus perspective for the study of sociolinguistic variation and suggests that some unexpected forms of socially structured variation may emerge if the analysis focuses on the basin of those rare productions that only very large databases may include.

3. Patterns of sociophonetic variation

The paper "Derhoticisation in Scottish English: A sociophonetic journey" by Jane Stuart-Smith, Eleanor Lawson, and James M. Scobbie presents sociophonetic data from the cities of the Central Belt of Scotland, Edinburgh and Glasgow, whose varieties show evidence of derhoticisation. According to Wells' (1982) taxonomy, Scottish English is usually thought to be a classic rhotic variety of English. Nevertheless, historical sources and several sociolinguistic inquiries have established that a derhoticisation process is present in selected varieties of Scottish English, though with variations related to the speakers' gender, social class, and speech style. The richness of points of view adopted in this study, ranging from auditory to acoustic and articulatory analyses, as well as from the discussion of different transcription methods to an investigation of the speakers' perceptual responses, forces the reader to reflect on what should be the best way of representing the complexity of sociophonetic data in our explanation of the speech processing mechanisms adopted by the speakers/hearers in normal linguistic interaction.

The link between the auditory, acoustic, and articulatory levels of analysis is a crucial point developed in the paper, inasmuch as the authors recognize that “each [level] gives a rather different (and incomplete) picture of the ‘same’ thing” (p. 68). The paper therefore combines the three different perspectives, including Ultrasound Tongue Imaging (UTI), in order to uncover the mechanisms of derhoticisation in production as well as perception. The question of whether and to what extent articulatory phonetics may (and has to) be integrated in traditional acoustics-based sociophonetic research is indeed a very controversial and topical issue (Celata & Calamai 2012). The paper by Stuart-Smith et al. offers a demonstration that it is possible to obtain natural, casual speech in an UTI experimental setting: according to the collected data, speech style appears to be more dependent on the speakers’ relationships with their interlocutors and the presence of friends and peers than with the experimental context in which data themselves are collected.

Particular attention is also devoted to the influence of broadcast media on language change. The results of a large-scale research project addressing the question of how London-based TV dramas exert their influence on Glaswegian vernacular phonology are summarized. The authors claim that the influence of the media language functions to emphasize the existent speech diversifications and accelerate sound changes in progress, which is more similar to internal developments than to the dynamics of language contact.

The listener is considered from two different points of view in the paper: as a phonetic analyst and as an “actor” parsing and responding to the different variants along the rhotic-derhotic continuum. In the first case, the problems of common practices of phonetic transcription is addressed. An experiment in which three expert phoneticians were requested to label different derhoticized variants illustrates that the transcribers mostly agreed on the number and the quality of the variants, while showing at the same time the existence of irreducible divergences concerning the position of the category boundaries. Concerning the second point, the reviewed studies show that derhoticisation also has a clear perceptual counterpart in the Glaswegian community. This is consistent with the view that both perception/imitation and production should be included in systematic socio-articulatory studies to develop a clearer picture of how articulatory variation spreads from speaker to hearer (e.g., Evans 2010) – a point that the authors emphasize repeatedly throughout the paper.

Rosalind Temple’s paper “Where and what is (t,d)? A case study in taking a step back in order to advance sociophonetics” departs from the following observation: in the sociophonetic literature, insufficient attention has been paid to the actual phonetic substance of some major variables, such as word-final coronal stop deletion in English, usually treated as a categorical variable rule. Word-final

coronal stop deletion represents indeed one of the most studied variables in English variationist sociolinguistics, and also one of the major focuses of the interaction between variationism and phonological theory (mostly from the point of view of Lexical Phonology). Nevertheless, according to the author, there are grounds for treating (t,d) as a function of common Connected Speech Processes observed by many phoneticians in English, rather than a particular variable rule restricted to word-final coronal stop deletion.

The aim of the study is therefore to demonstrate that similarities exist between the behaviour of word-final (t,d) stops and that of other word-final stop consonants. The forerunner of this approach can be found in the work of phoneticians (among whom, Francis Nolan, Paul Kerswill, and Susan Wright) who promoted the view that, in order to uncover the conditions on the occurrence of Connected Speech Processes, “it is necessary to adopt the techniques of sociolinguistics in conjunction with those of experimental phonetics” (Nolan & Kerswill 1989, p. 316). The paper by Temple integrally accepts this point of view by proposing an in-depth phonetic acoustic analysis of a large quantity of productions taken from the York Corpus of British English (Tagliamonte 1998), featuring a relatively standard variety of northern British English. The analysis shows that word-final (t,d) consonants “exhibit the same patterns of variability as other word-final stops” and “show parallel patterns of interaction with adjacent consonants resulting from Connected Speech Processes such as assimilation and cophonation” (p. 123).

The paper ends with a thorough discussion of different theoretical positions on the possibility of modeling the interaction of cognitive and physical phonetic effects to account for the observed phenomena of variation in naturalistic speech. The author provides arguments in support of those positions that tend to dismiss the idea of a sharp separation between cognitive and physiological constraints in phonetic effects, and recognizes that the development of articulatory phonetics and its adaptability to the dimensions and the requirements of sociophonetic research should be encouraged as it is expected to supply more direct evidence of the intricate interaction of cognitively and physiologically constrained effects in speech production.

In moving from Germanic to Romance languages it is necessary to reconsider the relationship between contemporary sociophonetics and traditional dialectology (or linguistic geography). The latter is probably to be viewed as the cultural root of sociolinguistic research in Europe. Yet the problematic nature of this heritage is widely acknowledged, at least to the extent that, on one side, “dialectology has been effectively isolated from general linguistics”, while on the other, “scholars continue to search for universal principles by manipulating isolated examples – subtracting from the available data, rather than adding to them” (Labov 1994:442). The Italo-Romance dialectal varieties, with their multidimensional repertoire of

uses, their geographic micro-diversifications and their extraordinary historical depth, are not only “the focus in the optical system of Romance linguistics” (Lausberg 1974: 252), but also an extremely attractive, thus far almost unexplored domain for sociophonetic excursions.

The papers by Giovanna Marotta and Rosanna Sornicola & Silvia Calamai offer two different sociophonetic rereadings of Italo-Romance dialectal phenomena.

G. Marotta’s “New parameters for the sociophonetic indexes. Evidence from the Tuscan varieties of Italian” is based on recent empirical work on the phonetic variability of Tuscan varieties. It aims at proposing a parametric evaluation of sociophonetic variation by making reference to the metaphor of solid bodies. Sociophonetic features can be viewed as solid bodies, i.e., entities that occupy a specific space in the domain of language and extend over a specific time span. They can be evaluated through a series of parameters, which correspond to the dimensions of the solid bodies – i.e., ‘Shape’, ‘Size’, ‘Thickness’, and ‘Weight’. These parameters summarize the distributional properties of a specific dialectal phenomenon with respect to its diffusion in the phonological system (e.g., the number of segments affected by a given phenomenon), across different speech styles (e.g., the degree of control that the speaker can exert over a certain pronunciation feature), and in the social community (e.g., the prestige or the stigma that a certain pronunciation feature may have within a given community). The parameters are also shown to be able to account for both categorical and optional or gradient properties of phonetic variation.

The examples are taken from the Tuscan dialectal repertoire: from Gorgia Toscana to l-velarization, from s-affrication to Rafforzamento Sintattico, and others. It is shown that the parameters are not independent of one another, at least in certain cases. This is due to the fact that they do not refer to the same level of linguistic description: ‘Shape’ and ‘Size’ are purely “descriptive parameters”, while “‘Thickness’ refers to the speaker’s behavior” and “‘Weight’ makes crucial reference to the listener” (p. 159). Therefore, a certain variation along the dimension of one parameter often carries the consequence of introducing a change in the value of a related parameter as well. For this reason, the author envisages among the future steps of the analysis the construction of a multi-factorial scale to account for the interdependence of selected parameters for individual phonetic phenomena of the Tuscan dialectal space, and the clarification of aspects of the “speaking-listening loop” that appears to be so crucial in the evaluation of socially structured variation in speech – as other papers of this volume equally emphasize.

The liveliness of Italo-Romance dialects and the importance of the sociophonetic values associated with local and regional features for the analysis of individual variation are also treated in Rosanna Sornicola & Silvia Calamai’s paper “Sound archives and linguistic variation: the case of the Phlegraean Diphthongs”.

The paper illustrates the usefulness of spontaneous speech sound archives for a better understanding of some crucial phonetic phenomena, such as spontaneous diphthongization, which are of interest in the domain of sociophonetics. As in several previous works (e.g., Sornicola 2002, 2006), the Campania regions assumes the role of a linguistic laboratory allowing the verification of different models of geographical, stylistic, and social variation. In this paper, particular attention is devoted to spontaneous speech analysis and the possibility of detecting the variability of the speaker's consciousness in adopting apparently contradictory speech behaviours, partly to be referred to local vernacular norms, partly to regional *koinés*, and partly to adherence to the standard language. The reflections of the author find their roots in Schuchardt, Jespersen, and Mathesius's thoughts and indirectly join the manifesto *Empirical foundations for a theory of Language Change* by Weinreich, Labov, Herzog (1968).

The Phlegrean area is characterized by considerable phenomena of diphthongization and vowel alteration which have only partially been studied by dialectologists. Phlegrean diphthongs represent an interesting case of "structural polymorphism", according to which the same structural unit appears in different forms because of segmental processes and/or the combined or alternate action of pragmatic and prosodic parameters. Such phenomena often happen beneath the speaker's level of awareness and appear to be highly irregular and unstable in diachronic terms. While former dialectologists such as Salvioni (1911) and Rohlfs (1949–54) only reported regular, unequivocal results (e.g. [e] > [ai], [o] > [au]), the meticulous examination of spontaneous, unplanned speech allows the sociophonetician to detect the highly variable nature of these diphthongs: the individual vowels follow different trajectories of diphthongisation inside the same text produced by the same speaker, according to variability in the prosodic and pragmatic conditions. The paper therefore demonstrates that micro-variationist analysis is an excellent tool for studying highly variable phenomena which seem to be indifferent to the traditional parameters of sociolinguistic variation (e.g. gender, age, social class).

The paper also stresses the potential of sound archives in offering phonetic data distributed over a long chronological stretch. One of the points that close the paper recalls Labov's (1994) reflections on the dichotomy between apparent time and real time: as linguists, we feel compelled to be able to trace linguistic changes over long periods of time. According to Labov (1994: 11), there are essentially two ways of accumulating real-time data: by "reviewing the past", and by "repeating the past". The limits of the first way are well known by field researchers: historical documents survive by chance and not by design, they are fragmentary and can only provide positive evidence. By contrast, to achieve the "repetition" of the past, it is necessary to return to the scene of a previous study and repeat it as closely as possible, in a time and money consuming field research ("it is important to

consider whether the outcome will be so decisive that the game is worth the candle"; Labov 1994: 75). In this respect, oral archives, digital preservation and audio restoration offer a substantial contribution to the study of language changes. The lack of phonetic records for instrumental measurements in the real time axis may be counterbalanced by a deeper exploration of this kind of Intangible Cultural Heritage, which is the result of a composite work of 'voice' preservation performed by dialectologists as well as anthropologists and ethnographers along the twentieth century (Ginouvé 2011). The exploration of sound archives may provide important insights for the development of a true historical experimental sociophonetics, as the Sornicola & Calamai's paper attempts at demonstrating.

4. Problematic sociophonetics

With Adrian Simpson's paper "Ejectives in English and German – Linguistic, sociophonetic, interactional, epiphenomenal?" the reader is engaged in a close inquiry on the apparent spread of ejectives in varieties of British English. The production of ejectives in English (as compared to German) is a case of articulatory variation that does not properly fit with the traditional classification of sociophonetic change.

Phonetic variation of English and German ejectives is analysed with respect to two different dimensions: function and production. The fine phonetic detail in ejectives articulation is discussed; the paper does not directly rely on a large and sociolinguistically stratified mass of speech data, as other papers of the volume do, though several hours of a television comedy are the reference data set of the analysis and attention is devoted to conversational situations. Moreover, it is argued that the analysis of the contexts in which the sound change unfolds must also proceed cautiously, as the structural context (e.g., word-finality) and the conversational context (e.g., spontaneous conversation with floor-holding pause) interact in natural language production in such a way that "within normal interaction, there are different categories of word-finality or pre-pausality, different contexts which may or may not be accompanied by different bundles of phonetic events" (p. 190).

Simpson's paper can therefore be seen as a purposely provocative conclusion to this volume, inasmuch as it clearly points out issues in the study of ongoing sound changes that are problematic for current sociophonetic research.

In particular, the sound change involving the apparent spread of ejectives in varieties of British English appears to be somewhat atypical for a number of reasons. First, it seems to have an internal and independent origin in several neighboring varieties, rather than a contact-induced source. Second, it is characterized by a rather low degree of predictability of occurrence in the different contexts,

since much variation is attested across speakers and also within the speech of individual speakers. Third, the auditory output is ambiguous with respect to the glottal and articulatory mechanisms, in such a way that the observed variability is inconsistent with some proposed theories of the propagation of sound changes from listeners' misinterpretations (e.g., Ohala 1974). The author appropriately argues that only a combination of different instrumental techniques, such as transillumination and airflow measurements, would produce a substantial advancement in the investigation of the articulatory mechanisms of ejectives' production – a position that reinforces similar arguments expressed in several other papers in the book. Fourth, the attested articulatory and functional ambiguity of ejectives is necessarily related to the very low number of occurrences in a given corpus. More specifically, the paper reports that in three hours of the television comedy *The Office*, only eight instances of ejectives were identified. This poses obvious problems from a methodological point of view: how would it be possible to apply a corpus perspective in the study of such infrequent speech phenomena? How should we track the precise development in time and space of marginal phonetic features by avoiding at the same time over-interpretations (besides misinterpretations) of the data? The author also points out that the level of phonetic-articulatory detail generally annotated in the available corpora is insufficient to the analysis of the spread of ejectives throughout the English language.

Some of these concerns, and particularly the latter, are probably also valid for several other phenomena worthy of sociophonetic investigation. This clearly encourages sociophonetics to pursue the journey through territories that in all probability still contain more surprises than might be expected.

5. Acknowledgments

The idea for this book originated during the international workshop “Sociophonetics, at the crossroads of speech variation, processing and communication”, which was held at Scuola Normale Superiore in Pisa in December 2010. The workshop attracted young linguists and consolidated experts from many corners of the globe, even unsuspected ones including South Africa and Japan, and appeared to inspire a lively debate on key concepts and new empirical evidence in the sociophonetic domain. A volume of proceedings with a selection of the papers that were presented at the conference came out in January 2012 (Calamai et al. 2012).

This book stemmed from the resolution of making the most pertinent concepts and hypotheses presented at the workshop available to a larger audience. We therefore asked several of our invited speakers to supply material that would

contribute to a comprehensive overview of the issues tackled in contemporary sociophonetic research.

We are grateful to all contributors for their valuable collaboration. Most of the chapters were patiently rewritten more than once in order to improve the book's internal consistency and readability. We gratefully acknowledge the editors of the Studies in Language Variation series for their careful comments and suggestions throughout the various stages of the work. We are particularly grateful to Pier Marco Bertinetto for having provided us with the unique opportunity of organizing the sociophonetic workshop in Pisa, and for having supported all our subsequent scientific and editorial initiatives.

Our hope is that the volume will be a stimulus to further productive inquiry into the nature of sociophonetic variation and the way in which the speakers and hearers of a language organize sociophonetic information in their mental representation of speech.

References

- Baayen, Harald R., Robert Schreuder, Nivja H. De Jong & Andrea Krott. 2002. "Dutch inflection: the rules that prove the exception". *Storage and Computation in the Language Faculty* ed. by Sieb Nootboom, Fred Weerman & Franz Wijnen, 61–92. Dordrecht: Kluwer.
DOI: 10.1007/978-94-010-0355-1_3
- Bybee, Joan L. 2000. "The phonology of the lexicon: Evidence from lexical diffusion". *Usage-based models of language* ed. by Michael Barlow & Suzanne Kemmer, 65–85. Stanford: CSLI Publications.
- Bybee, Joan L. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.
DOI: 10.1017/CBO9780511612886
- Bybee, Joan L. 2005. "La liaison: effets de fréquence et constructions". *Langages* 158. 24–37.
DOI: 10.3917/lang.158.0024
- Calamai, Silvia, Chiara Celata & Luca Ciucci, eds. 2012. *Proceedings of "Sociophonetics, at the crossroads of speech variation, processing and communication" (Pisa, December 14–15, 2010)*. Pisa, Edizioni della Normale. <http://www.sns.it/scuola/edizioni/testonline/>
- Celata, Chiara & Silvia Calamai, eds. 2012. *Articulatory techniques for sociophonetic research*. (= *Italian Journal of Linguistics*, 24/1). Pisa: Pacini.
- Donegan, Patricia & David Stampe. 1979. "The study of Natural Phonology". *Current Approaches to Phonological Theory* ed. by Daniel A. Dinnsen, 126–173. Bloomington: Indiana University Press.
- Durand, Jacques, Bernard Laks & Chantal Lyche. 2002. "La phonologie du français contemporain: usages, variétés et structures". *Romanistische Korpuslinguistik: Korpora und gesprochene Sprache / Romance Corpus Linguistics: Corpora and Spoken Language* ed. by Claus D. Pusch & Wolfgang Raible, 93–106. Tübingen: Narr.
- Durand, Jacques, Bernard Laks, Basilio Calderone & Atanas Tchobanov. 2011. "Que savons-nous de la liaison aujourd'hui?". *Langue Française* 169. 103–135.
DOI: 10.3917/lf.169.0103

- Evans, Betsy E. 2010. "Aspects of the acoustic study of imitation". *A reader in sociophonetics* ed. by Dennis R. Preston & Nancy Niedzielski, 379–391. New York: Mouton De Gruyter.
- Foulkes, Paul. 2010. "Exploring social-indexical variation: a long past but a short history". *Laboratory Phonology* 1. 5–39. DOI: 10.1515/labphon.2010.003
- Ginouès, Véronique. 2011. "Quand le renard raconte ses histoires au monde. La naissance du portail du patrimoine orale, catalogue collectif d'archives sonores et audiovisuelles". *Le patrimoine culturel immatériel: premières expériences en France* (= *Internationale de l'Imaginaire* 25) ed. by Christian Hottin 107–128. <http://halshs.archives-ouvertes.fr/halshs-00588487/fr/>
- Goldinger, Stephen D. 1998. "Echoes of echoes? An episodic theory of lexical access". *Psychological Review* 105. 251–279. DOI: 10.1037/0033-295X.105.2.251
- Jannedy, Stefanie & Jennifer Hay. 2006. "Editorial. Modelling Sociophonetic Variation". *Journal of Phonetics* 34/4. 405–408. DOI: 10.1016/j.wocn.2006.08.001
- Köpke, Barbara. 2004. "Neurolinguistic aspects of attrition". *Journal of Neurolinguistics* 17. 3–30. DOI: 10.1016/S0911-6044(03)00051-4
- Labov, William. 1994. *Principles of Linguistic Change: Linguistic Factors*. Oxford: Blackwell.
- Lausberg, Heinrich. 1974. *Noterelle di dialettologia italiana*. Göttingen: Vandenhoeck & Ruprecht Verlag.
- Nolan, Francis & Paul Kerswill. 1989. "The description of connected speech processes". *An introduction to the pronunciation of English* ed. by Alfred C. Gimson, 4th edition revised by S. Ramsaran, 295–316. London: Arnold.
- Ohala, John J. 1974. "Experimental historical phonology". *Historical linguistics II. Theory and description in phonology. Proceedings of the 1st International Conference on Historical Linguistics (Edinburgh, September 2–7, 1973)* ed. by John M. Anderson & Charles Jones, 353–389. Amsterdam: North Holland.
- Pierrehumbert, Janet. 2001. "Exemplar dynamics: Word frequency, lenition and contrast". *Frequency and the emergence of linguistic structure* ed. by Joan L. Bybee & Paul Hopper, 137–158. Amsterdam & Philadelphia: John Benjamins.
- Rohlf, Gerhard. 1949–1954. *Historische Grammatik der Italienischen Sprache und ihrer Mundarten*. Bern: Francke.
- Salvioni, Carlo. 1911. "Zur Lautgeschichte. Appunti per la storia del vocalismo tonico italiano". *Zeitschrift für Romanische Philologie* 35. 486–488.
- Sornicola, Rosanna. 2002. "La variazione dialettale nell'area costiera napoletana. Il progetto di un archivio di testi dialettali parlati". *Bollettino Linguistico Campano* 1. 131–155.
- Sornicola, Rosanna. 2006. "Dialectology and history. The problem of the Adriatic-Tyrrhenian dialect corridor". *Rethinking Languages in Contact. The Case of Italian* ed. by Anna Laura Lepschy, Giulio C. Lepschy & Arturo Tosi, 127–145. Oxford: Legenda.
- Tagliamonte, Sali A. 1998. "Was/were variation across the generations: View from the city of York". *Language Variation and Change* 10. 153–91. DOI: 10.1017/S0954394500001277
- Weinreich, Uriel, William Labov & Marvin I. Herzog. 1968. "Empirical foundations for a theory of language change". *Directions for Historical Linguistics: a Symposium* ed. Winfred P. Lehmann & Yakov Malkiel, 95–188. Austin: University of Texas Press.
- Wells, John C. 1982. *Accents of English*. Cambridge: Cambridge University Press.

PART I

Variation and sociolinguistics

The sociophonetic orientation of the language learner

William Labov

University of Pennsylvania

This paper is an effort to define the phonetic target of the language learner: what are the data that the child focuses on in becoming a native speaker? A number of studies are reviewed to show that children reject the idiosyncratic features of their parents' phonetic system if they do not match the pattern of the larger speech community: in the acquisition of the Philadelphia and New York City dialects; the formation of a new dialect in Milton Keynes; the spread of the low back merger in eastern New England; the reduction of the future marker in Tok Pisin. The end result is a high degree of uniformity in both the categorical and variable aspects of language, where individual variation is reduced below the level of linguistic significance.

1. Introduction

This paper is an attempt to define the target of the child who is engaged in acquiring the phonetics and phonology of a language: asking, what are the data that the child attends to in the process of becoming a native speaker? The argument to be advanced here is that the human language learning capacity is aimed at the acquisition of the most general community pattern. The end result is a high degree of uniformity in both the categorical and variable aspects of language production, where individual variation is reduced below the level of linguistic significance. This approach to the nature of language is aligned to the *central dogma of sociolinguistics*:

1. the community is conceptually and analytically prior to the individual.

For linguistic analysis, this means that the behavior of an individual can be understood only through the study of the social groups of which he or she is a member. Following the approach outlined in Weinreich et al. (1968), language is seen as an abstract pattern located in the speech community and exterior to the individual.

Language is a *social fact*. In Durkheim's terms, "ways of behaving, thinking and feeling, exterior to the individual, which possess a power of coercion by which they are imposed on him" (Durkheim 1895:5 [my translation]).

The human language faculty, an evolutionary development rooted in human physiology, is then viewed as the capacity to perceive, reproduce and employ such generalized patterns.

The opposing point of view, common even among students of the speech community, is that the individual constructs a grammar on the basis of the particular set of input data to which he or she is exposed in the formative years. Since in this view, the language learning mechanism is not programmed to delete idiosyncratic constructs, the end result is that each learner winds up with a particular version of the grammar based on individual experience. The speech community is then seen as a vague average or assembly of these idiolectal variants.

Enthusiasm for the individual is not a new development. Thus Durkheim notes, "the word *coercion*, by which we define [social facts], has a risk of irritating the zealous partisans of an absolute individualism. As they believe that the individual is perfectly autonomous, they feel that the individual is diminished each time that it seems that he does not act entirely by himself" (Durkheim 1895:6).

Although the critique of the idiolect in Weinreich et al. (1968) was widely accepted, a tendency to focus on the individual recurs with striking regularity. So Janet Holmes on *Sociolinguistics and the Individual* writes

The linguist should be able to pin-point the development of a language as a result of individual choices, and [...] the sociolinguist should try to relate changes in social structure to changes in individual cultural values as expressed through speech in social interaction. Individual behavior is thus seen as the proper starting point for sociolinguistic investigation. (Holmes 1969)

Le Page & Tabouret-Keller (1985) see language as essentially idiosyncratic. Language is for them the linguistic repertoire of the individual; the individual is "the locus of his language" (Le Page & Tabouret-Keller 1985:116). Johnstone's book on *The Linguistic Individual* (1996) is devoted to the argument that we should think about language from the perspective of the individual speaker, rather than the perspective of the social aggregate or the abstract linguistic system.

The general perspective put forward here reinforces the contrary view. It is argued that the individual does not exist as a unit of linguistic analysis. Though the recordings and judgments of sociolinguistic research are gathered from individual speakers, their idiosyncratic behavior is not our focus, but rather the extent to which they respond to wider community patterns.

The compulsion for the language learner to turn outward may be thought of as the end result of a competition between two types of language learning. In one

type each individual forms a grammar that is informed primarily by the initial input and is the unique product of his or her individual experience. The other type is programmed to keep smoothing the result with data from following contacts (up to a certain limit). This second type is *outwardly bound*, in the sense that it searches for the community pattern as epistemologically prior to the individual pattern.

The suggestion here is that from an evolutionary point of view, the second type will survive, and has survived, at the expense of the first.

Most readers will have personal evidence that children do not adopt those features of their parents' dialect that fail to match the pattern of the surrounding community. Linguists are especially conscious of such mismatches with their parents or with their children. For my general argument, I will want to go beyond this personal experience and consider systematic studies of this process.

2. Rejection of parental idiosyncrasy

2.1 The King of Prussia study

A clear view of how children reject their parents' dialect can be drawn from that part of the Philadelphia Neighborhood Study which was designed by Payne (1976, 1980). For the upper middle class neighborhood, she selected King of Prussia, a new suburb that barely existed in the 1940s. A rapid development of electronic and chemical industries drew half of the population from metropolitan Philadelphia and half from out-of-state communities with very different vowel systems: Massachusetts, New York, and Cleveland.

Payne's study included the acquisition of the Philadelphia dialect by 34 children of out-of-state parents. In Figure 1, the vertical axis is the percent of children who consistently rejected their parents' dialect in favor of the Philadelphia pattern. From left to right we have the fronting of (aw) to [ɛo] or [eɔ], the centralization of /ay/ before voiceless finals in *like*, *right*, *fight* before voiceless consonants, the fronting of (ow) in *go*, *boat*, *road*, the raising of the nucleus of (oy) in *choice*, *boy*, and the fronting of (uw) in *do*, *dew*, *move*, etc.

The two upper lines show that the majority of children who spent at least half of their formative years in Philadelphia departed from their parents' pattern consistently – those who arrived from birth to 4 and from 5 to 9 years old. Those who came later did not, except for the fronting of /ow/. But conversely it should be noted that a third at least still showed some traces of the parental system. Though parents are not the target of language learning, they are not without influence.

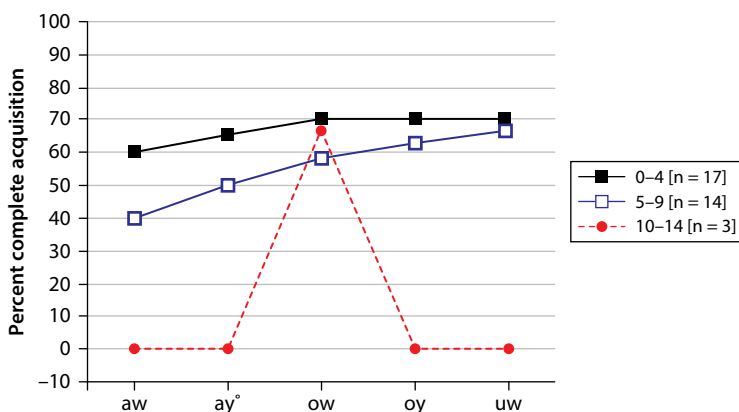


Figure 1. Acquisition of Philadelphia variables by 34 children of out-of-state families in King of Prussia by age of arrival. Based on Payne (1980).

2.2 Milton Keynes

A second study of a new town was Kerswill and Williams' research in Milton Keynes, composed of immigrants from many different areas of England (Williams & Kerswill 1999; Kerswill & Williams 2000). Milton Keynes did not exist in 1971, but grew to 176,000 in 1991. Three quarters of its residents came from the southeastern England: 35% from London, 32% from other southern counties, but only 3% from the immediate sub-region within 15 minutes drive. This project was elegantly and carefully designed to record the phonology of 8 boys and 8 girls at each of three age levels: 4, 8 and 12 years old, together with the caretakers of each, a total of 96 speakers. The new Milton Keynes dialect that arose was a distinct entity to combining the some features of London and the home counties with some remnants of the local dialect.

Kerswill and Williams provide Figure 2 for the Milton Keynes development of the variable (ow), in *goat*, *go*, *road*, etc. The distribution of phonetic variants are shown for the three age levels of children together with that for their female caretakers. The horizontal axis shows the frequencies of three variants: front nuclei with unrounded glides, front nuclei with rounded glides, and back upgliding vowels with a central or back glide. The four-year-olds have clearly not departed from the pattern of their parents and caretakers following them in a moderate percent of the second variant and a high representation of the third. But the eight and twelve-year-olds have, shifting massively to the front nuclei with rounded glides. Though Kerswill and Williams do not match individual children with their caretakers, it is clear that by eight years of age, these children no longer take their parents' vowel systems as the target for language learning.

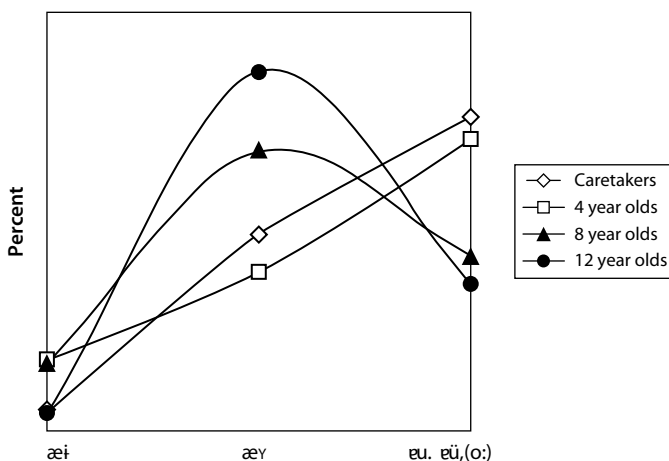


Figure 2. Development of local phonetic forms for (ow) in *goat*, *go*, *road* etc. in Milton Keynes by age. From Kerswill and Williams (1994).

2.3 The future in Tok Pisin

A dramatic view of the disconnect between parents and children is found when a pidgin language, with no native speakers shifts to a creole as a generation of children grow up with this as their first language. Bickerton (1981) treats these as linguistically orphaned, forced to use their naked language learning ability because they do not recognize their parents speech as legitimate language. Sankoff & Laberge (1973) studied this process in the development of the future marker BAI of Tok Pisin (derived from *baimbai*). Adults tended to give this formative secondary stress, realizing 'He will go' as /em bai i-go/, while children tended to reduce /bai/ to [bə] in [embəigo] or even [embigo].

(1) Parent	Child
em bai i-go	em b(ə) i-go

Figure 3 shows the maintenance of secondary stress of *bai* on the vertical axis with age on the horizontal axis. As a whole, parents are very different from their children: 35 to 70% *bai* with secondary stress and children at much lower levels: 5 to 45%. However, when Sankoff added to her original diagram the lines connecting parents to children, we observe that they are largely parallel. As the King of Prussia study in §2.1 showed, the influence of the parental input is not completely eliminated even when children have absorbed the new community norm.

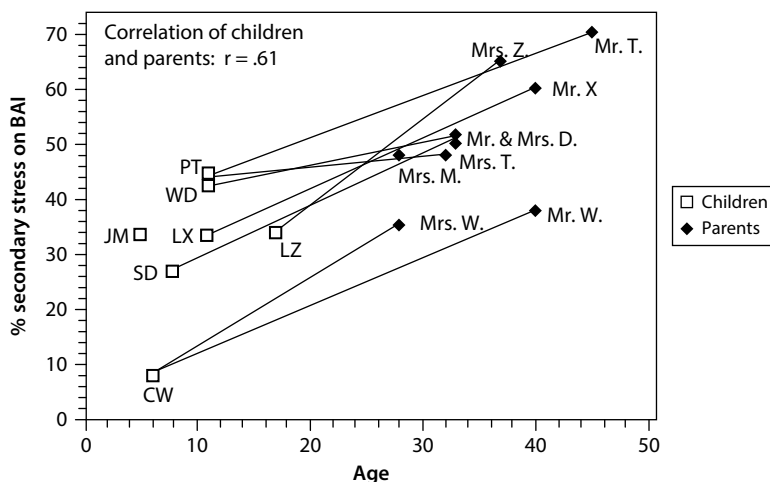


Figure 3. Percent secondary stress on future marker BAI of Tok Pisin by parents and children (source: Sankoff & Laberge 1973 and G. Sankoff, personal communication).

2.4 The low back merger in Eastern Massachusetts

Children's acquisition of a new change can also be observed quite clearly in Johnson's study of the spread of the low back merger in eastern Massachusetts (2009). In the area of eastern Massachusetts around Boston, there is a complete merger of /o/ and /oh/ in *cot* and *caught*, *Otto* and *auto*, etc. while in the region next to Rhode Island, this distinction is maintained.¹ Figure 4 is Johnson's map of the boundary that he established between the merged region on the right and the unmerged region on the left. He found that this boundary was stable across two or three adult generations but when he studied families in the border towns of Attleboro and Seekonk, he found a rapid shift towards the merger in the youngest generation.

Figure 5 shows the state of the merger among the children of the families that Johnson studied in the town of Seekonk. The vertical axis registers Seekonk children's grade in school, and along the horizontal axis are grouped the various families according to whether mothers, fathers, or both make the distinction. In grades 2–6 there is a belt of 10 black symbols indicating complete merger, while older children maintain the distinct system of the community and younger children have not completely emerged from the influence of their parents.

1. As the captions indicate, the area to the west where /oh/ and /o/ are distinct shows the merger of /ah/ in *father* with /o/ in *bother*, while these are distinct in the area to the east where /o/ and /oh/ are merged.

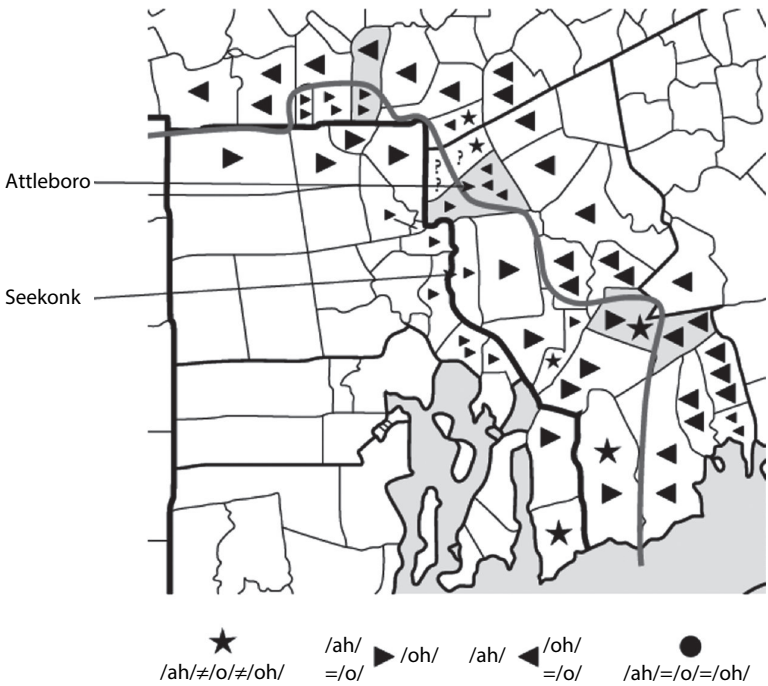


Figure 4. Western boundary of low back merger area in eastern Massachusetts established by Johnson (2009).

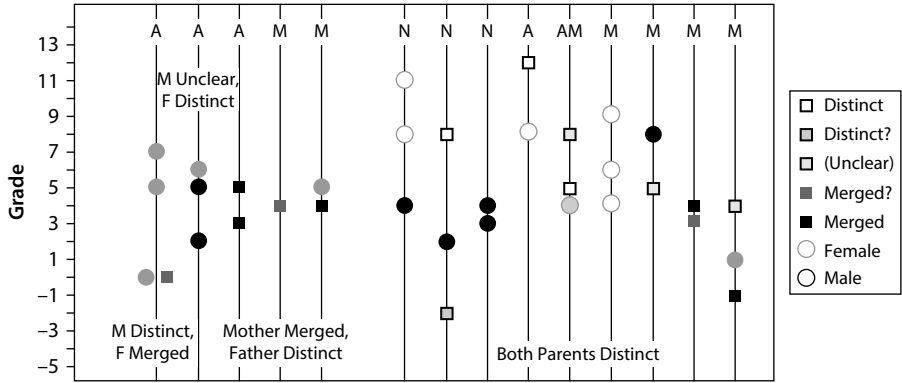


Figure 5. Low back vowel systems in Seekonk children by grade and parental system. Elementary schools: A = Aitkin, M = Martin, N = North (source: Johnson 2010, Figure 5.3).

But although parents' influence may be eroded, it is not eliminated immediately as children mature. We can trace the persistence of the waning influence of parents by comparing the relative influence of mothers and fathers. Since female caretakers are the primary source for first language learning, we can expect the effect of having a mother with a distinct system would outweigh the effect of a father with a distinct system. Table 1 is drawn from Johnson's School Survey in Attleboro. The figures register a distinctness score where 4 is maximum in children's response to minimal pairs as 'same' or 'different'. We see that having a mother who makes the distinction (but a father who doesn't) yields twice as high a distinctness score as the reverse situation.

Table 1. Distinctness scores for Attleboro 8th grade subjects cross tabulated by parents' dialect origins (Johnson 2010). 4 = clearly distinct in two minimal pairs. 1 = clearly merged in two minimal pairs.

	Distinct mother	Merged mother
Distinct father	2.59 (N = 24)	0.83 (N = 6)
Merged father	1.67 (N = 6)	0.70 (N = 37)

2.5 The change of apical to uvular /r/ in Montreal

We can now shift to a view of the rejection of parental influence by a slightly older group: adolescents. Figure 6 is from Sankoff and Blondeau's (2007) real-time re-study of the shift from apical to uvular /r/ in Montreal. The traditional Quebecois dialect of Montreal used consistent apical /r/ in *rouge*, *arrive*, *pour*, etc. The new uvular form spread geographically from the Quebec City area, though it was also identified with continental standard French.

In Figure 6 the percent uvular [R] appears on the vertical axis and age of the speaker on the horizontal axis. Black symbols show percent [R] in 1971 and the grey symbols in 1984. The four black arrows connect the 1971 and 1984 values of speakers who increased their use of the uvular variant from some intermediate value in 1971. These four adolescents had not departed completely from their parents in 1971, but 13 years later, in 1984, they had done so.

At upper left are seven black diamonds representing adolescents who had already achieved 100% uvular (R) by 1971 and maintained it in 1984. The parents of these seven adolescents were part of a community that used 100% apical (r). In fact, when parents of these adolescents do appear in the recordings of the 1971 interviews, they do use 100% apical [r]. We conclude that 100% apical would have been their model in early L1 acquisition. It is clear that the majority of adolescent speakers in 1971 had taken as their target a form of /r/ quite distinct from that of their parents and 7 out of 12 achieved consistent control over this uvular form.

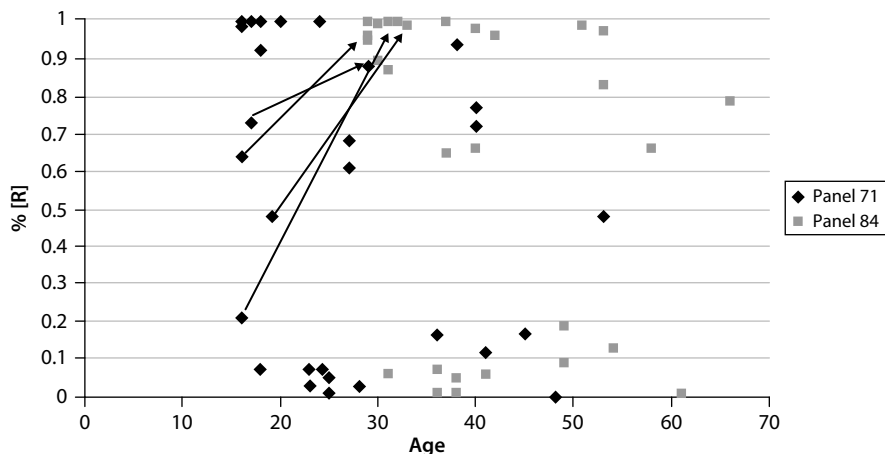


Figure 6. Shift of /r/ from apical [r] to uvular [R] in Montreal (from Sankoff and Blondeau 2007).

2.6 Second vs. third generation in New York City

The pattern seen with relative influence of competing dialects is repeated with even greater clarity when the caretakers are not native speakers of the dominant language of the speech community. With the massive immigration into the U.S. from southern and eastern Europe in the late 19th century, large numbers of children acquired English in households where the caretakers were native speakers of Italian, Polish, Yiddish, Greek, and many other languages. There is ample evidence that the foreign accent of the parent has no influence on the adult English of the next generation.

This appears most clearly in a comparison of second and third generations that I was able to make in the 1963 study of New Yorkers on the Lower East Side (Labov 1966). In the sample of 81 adults, it was possible to make two such comparisons. Among nine upper middle class Jewish men, age 21–39 years, three were 2nd generation, with Yiddish-speaking parents, and six were third generation, with English speaking parents. Among nine working class Jewish women, age 40–65, six were second generation and three were third generation. The linguistic variables were (æh) and (oh). The raising of (æh) carries [bæd] to [be:əd] and [bi:əd], while the raising of /oh/ carries [kɔfi] to [kuəfi]. Both use a 40-point scale of height: for (oh), 10 shows a consistent low lax vowel close to [ʊ] and 40 shows consistent high tense ingliding vowels approximating [uə]. Both of these sound changes run counter to a direct influence of Yiddish on English, which gives us lax [ə bəd mən] instead of tense [be:əd me:ən] and [ə kʊp kʊfi] instead of [ə kʌp ku:əfi]. Table 2 shows the results of this comparison.

Table 2. Comparison of two vowel variables in second and third generations in New York City.

	Upper middle class Jewish				Working class Jewish			
	Younger men [21–30 years]				Older women [40–65 years]			
	/æh/ scores		/oh/ scores		/æh/ scores		/oh/ scores	
	2nd gen	3rd gen	2nd gen	3rd gen	2nd gen	3rd gen	2nd gen	3rd gen
N	3	6	3	6	6	3	6	3
Mean	31.0	30.3	27.0	24.5	28.6	28.6	19.9	20.6
Std dev	7.8	5.8	1.6	4.8	3.9	10.9	3.5	1.1
t-test	0.15		0.86		0.47		0.33	

There are no significant differences between generations for either group in the raising of (æh). Nor are there any significant differences between the 2nd and 3rd generations in the raising of (oh). We see that the eventual result for young and middle aged adults is the same for children whose parents were speakers of Yiddish and those whose parents were native speakers of the New York City dialect.

2.7 The effect of ethnicity on sound change in Philadelphia

In the 1970s, the Philadelphia Neighborhood Study created a stratified sample of the white mainstream areas of Philadelphia in ten neighborhoods that embodied a full range of social classes and ethnicities (Labov 1980, 2001). The circles on Figure 7 shows the mean values of the vowels of 116 speakers. The arrows indicate the direction of change as determined by the age coefficients of step-wise multiple regression. The head of the arrow represents the expected values for speakers 25 years younger than the mean age for the sample; the tails of the arrow show expected values for speakers 25 years older than the mean.² The three largest arrows indicate the new and vigorous changes in the system; here we are focusing on the raising and fronting of /ey/ in checked syllables: *main*, *paid*, *late*, etc. This change in progress was first recognized in the course of the acoustic analysis of the 116 speakers but is found consistently across all social groups. These results from apparent time studies were confirmed by a re-study in real time by Conn (2005).

Table 3 shows the full output of the regression analysis for the fronting of /ey/ in checked syllables. Age is negative and significant at < .0001 probability, as Figure 7 indicated: that is, the younger the speaker, the higher the value of the

2. Expected values for F1 and F2 are calculated by multiplying the age coefficients in the output of the regression times 25 or –25 and adding this to the regression constant.

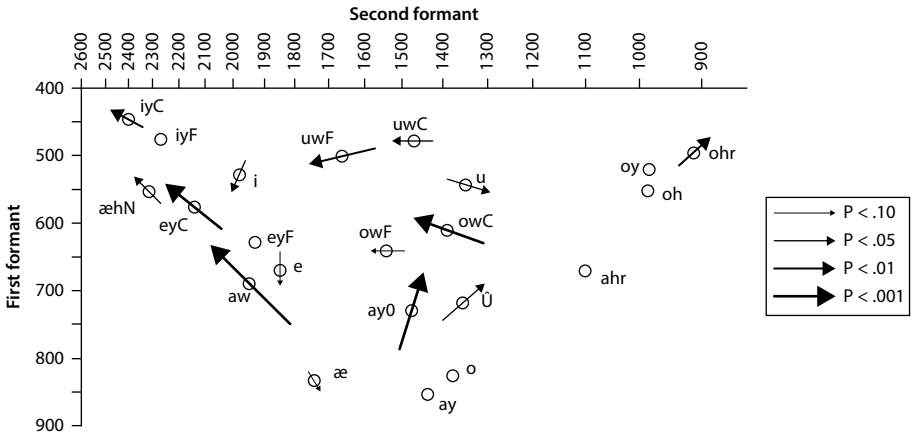


Figure 7. Mean values of vowels for the Philadelphia Neighborhood Study (N = 116). Arrows represent age coefficients in stepwise regression. Head of arrow: expected value for speakers 25 years younger than the mean; tail or arrow expected value for speakers 25 years older than the mean.

second formant. Female gender is positive at an amount equivalent to a 25 year difference in age at $< .01$ probability. Membership in the upper working class (compared to the residual group, the lower working class) is a somewhat larger effect, at $< .05$ probability level. Residence in the oldest settled working class neighborhood (“Wicket St.” in Kensington) contributes even more, again $< .01$. But Jewish ethnicity is not a significant effect, nor is Italian, Irish, WASP or German. Finally, I entered Generational status, to ensure comparison with the New York City data: this too fails to show a significant effect. In other words, the use of a second language or dialect has no linguistic consequence for those who follow.

Table 3. Stepwise regression output for fronting of checked /ey/ in the Philadelphia Neighborhood Study [N = 112] as registered by second formant measurements.

Variable	Coefficient	Probability
Age (* 25 yrs)	-85	≤ 0.0001
Female	83	0.008
Upper working class	108	0.026
Wicket St. neighborhood	145	0.004
Jewish	-169	n.s.
Italian	38	n.s.
Irish	-2	n.s.
Wasp	-91	n.s.
German	-98	n.s.
Generational status	9	n.s.

3. Where ethnicity emerges

Some recent studies indicate situations where ethnicity and foreign language background do emerge as significant factors in community studies. Boberg (2004) finds strong differentiation of Irish, Italian and Jewish speakers of English in Montreal. He attributes this to the minority status of English, where distinct ethnic neighborhoods are separated by French-speaking areas and exposure to general Canadian English is limited. Wagner (2013) reports small differences between Irish- and Italian-American girls in the contact in a Philadelphia high school. Here the ethnic effect appears to be mediated by the fact that the Irish group favored the stance of “toughness” associated with linguistic features like the backing of centralized (ay0) in *light*, *fight*, etc. These linguistic effects of ethnicity do not run counter to the major theme of this article: the fact that children reject those features of their parents’ language that deviate from the speech community in which they are raised.

4. Conclusion

We then have gained some idea of what is to be learned, and how the language learner looks outward to master the broader community patterns. The study of linguistic variation is sometimes pursued as a way of showing how different people are from one another. And it is perfectly true that the larger our data base becomes, the more often will statistical analysis reveal subtle differences among subgroups of the population. I have tried to turn the focus away from those minor subdivisions and ask us to account for the uniformities that result from the outward orientation of the language learning faculty. Though we have much to learn from micro-analysis, we have more to learn from our efforts to grasp the larger pattern.

References

- Bickerton, Derek. 1981. *Roots of language*. Ann Arbor: Karoma Publishers.
- Boberg, Charles. 2004. “Ethnic patterns in the phonetics of Montreal English”. *Journal of Sociolinguistics* 8. 538–568.
- Durkheim, Emile. 1895. *Les Règles de la Méthode Sociologique*. Paris: Librairie Félix Alcan.
- Johnson, Daniel Ezra. 2009. “Stability and change across a dialect boundary: the low vowels of southeastern New England”. *Publications of the American Dialect Society* 95.
- Kerswill, Paul & Anne Williams. 2000. “Children, adolescents and language change”. *Language Variation and Change* 8. 177–202. DOI: 10.1017/S0954394500001137
- Holmes, Janet. 1969. “Sociolinguistics and the individual”. *Tè Reo* 12. 41–47.

- Labov, William. 1980. "The social origins of sound change". *Locating Language in Time and Space* ed. by William Labov, 251–266. New York: Academic Press.
- Labov, William. 2001. *Principles of Linguistic Change. Volume 2: Social factors*. Oxford: Blackwell.
- Labov, William. 1966. *The Social Stratification of English in New York City*. Washington: Center for Applied Linguistics. (2nd edition: Cambridge, Cambridge University Press. 2006.) DOI: 10.1017/CBO9780511618208
- Le Page, Robert B. & Andree Tabouret-Keller. 1985. *Acts of Identity: Creole-based approaches to language and ethnicity*. Cambridge: Cambridge University Press.
- Payne, Arvilla. 1976. *The acquisition of the phonological system of a second dialect*. PhD Dissertation, University of Pennsylvania.
- Payne, Arvilla. 1980. "Factors controlling the acquisition of the Philadelphia dialect by out-of-state children". *Locating Language in Time and Space* ed. by William Labov, 143–178. New York: Academic Press.
- Sankoff, Gillian & Hélène Blondeau. 2007. "Language change across the lifespan: /r/ in Montreal French". *Language* 83. 560–588. DOI: 10.1353/lan.2007.0106
- Sankoff, Gillian & Suzanne Laberge. 1973. "On the Acquisition of Native Speakers by a Language". *Kivung* 6. 32–47.
- Sober, Elliott. 1994. *Conceptual Issues in Evolutionary Biology*. Cambridge: MIT Press.
- Wagner, Suzanne Evans. 2013. "'We act like girls and we don't act like men.' Ethnicity and local language in a Philadelphia high school". *Language in Society* 42. 361–383.
- Weinreich, Uriel, William Labov & Marvin Herzog. 1968. "Empirical foundations for a theory of language change". *Directions for Historical Linguistics* ed. by Winfred Lehmann & Yakov Malkiel, 97–195. Austin: University of Texas Press.
- Williams, Ann & Paul Kerswill. 1999. "Dialect levelling: continuity vs. change in Milton Keynes, Reading and Hull". *Urban Voices* ed. by Paul Foulkes & Gerard Docherty, 141–162. London: Arnold.

French liaison and the lexical repository

Bernard Laks*, Basilio Calderone** and Chiara Celata***

* Université de Paris Ouest Nanterre La Défense,
Institut Universitaire de France

** CLLE-ERSS, CNRS & Université de Toulouse – Le Mirail

*** Scuola Normale Superiore di Pisa

In this paper we propose a frequency analysis of French liaison that focuses on the liaison environments attested in the PFC database. The results of the analysis show the existence of a significant relationship (statistically interpreted as a power-law distribution) according to which a very restricted set of liaison environments has very high frequency of occurrence in the corpus and is substantially untouched by phonological and sociolinguistic variation, while a large “periphery” of infrequent uses appears to show significant aspects of style- and speaker-dependent variation. The study therefore demonstrates the importance of basing any variationist analysis on very large data sample, such as those provided by contemporary, well-reasoned linguistic corpora.

1. Introduction: *Datum* and *exemplum* approaches in the study of phonological variation

That the phonology of a living language should be based on the description and analysis of actual usage, as manifested by the concrete occurrences in the language in question, is a proposition which may seem to be self-evident. In the contemporary period, however, scientific practice does not support it entirely, and the use of corpora appears to be a minority option among many research orientations (Laks 2008, 2011).

With the Chomskyan critique of the finite nature of corpora and the limits of the syntagmatic model (Chomsky 1957, 1965), the paradigm of the *exemplum* (Laks 2008) was to dominate the field for a period of more than thirty years. Whatever the question to be addressed, it sufficed to invoke a small number of examples deemed to be pertinent, or even crucial for a given reasoning, to support specific hypotheses as well as broad theoretical assumptions. The phonology of French did not remain untouched by the generative drift of the period. Even before the publication of SPE,

the first doctoral thesis applying the framework and methods of SPE was defended. Schane (1965) thus devoted a 45 page chapter to the analysis of French *schwa*, consonantal liaison, *h aspiré*, liaisons within the verbal group, inversions and final fixed consonants. Notwithstanding the complexity of the phenomena addressed, which had hitherto produced thousands of pages of analysis, Schane's very modest chapter based treatment of those six issues in French phonology on a total of 73 examples. In all, the thesis put forward a total of 41 ordered rules. It was up to the reader to assess such a ratio between rules and examples.¹

During the 20th century the surge in the power of linguistic corpora and databases, coupled with the development of efficient sampling tools in the framework of computational linguistics, has facilitated the availability of an ever-larger and more varied range of linguistic data. Sociolinguistics (especially 'Labovian' variationism) had a great role in developing the paradigm of *datum* in linguistics (Labov 2004). The problem of structured heterogeneity and inherent variation in the grammar of any language system has been the focus of the scientific approach since Weinreich et al. (1968); sociolinguistics presents itself as a corpus linguistics that takes seriously the internal social organization of the linguistic community being examined. However, with the contemporary quantitative approaches and the development of corpus linguistics proper the use of the *datum* is no longer limited to the description of social variation; on the contrary, it reaches to the very heart of linguistic explanations in several domains such as acquisition, lexical storage, community patterns of variation, heterogeneous linguistic competence (such as pidginization, creolization, dialect formation or loss). The scientific study of linguistic usage (Langacker 2000; Bybee 2006) and corpus linguistics converge in these contemporary approaches, giving rise to quantitative and variable models of language acquisition in psycholinguistics (e.g., Tomasello 2003, 2008), probabilistic and stochastic models of language variation in formal descriptions (e.g., Boersma & Hamann 2009), and contemporary sociophonology, that very strongly shows the traces of this new empiricism in linguistic studies. At the crossroads of traditional sociolinguistics and experimental phonology, contemporary sociophonology investigates the phenomena of speech use and comprehension from a cognitivist perspective and explicitly borrows themes and concepts from experimental psychology, particularly from exemplar theory and the idea that multi-sensorial mnemonic traces are stored in the mind of the speaker according to their distribution (frequency) and recoverability (recency) (Goldinger 1998; Foulkes & Docherty 2006; Johnson 2006). In recent sociophonological studies, the emergence of a linguistic category (be it phonological or socio-indexical) is predicted on the basis

1. Twenty-six examples for truncation and elision, 14 for fixed final consonants and numeral adjectives, 3 for aspirated h, 2 for hiatus, 20 for junctures in verbal groups and for inversions, 8 for postpositioned pronouns. Twenty-two additional examples are quoted in the footnotes.

of observable mechanisms of statistically oriented learning and the formation of secondary assemblages within the classes of exemplars (Foulkes 2006). Such predictability is therefore based on the acknowledgement of corpus linguistics as a normal sociophonological practice (see also the Sornicola's paper in this volume).

It was within this theoretical and empirical context, characterised by the return to the forefront of a corpus-based linguistics, that the "Phonologie du Français Contemporain" (PFC) programme has been constructed and developed since 1999 (Durand et al. 2002, 2005). The aim of the PFC programme is to construct a significant repository of contemporary French that will enable researchers to address the diversity of the oral usages of the language, both within France and in the wider French-speaking world.² The construction of a large database which has been devised, labelled and standardised in order to allow as many types of secondary analyses as possible, rests at the heart of the programme whose initial objectives can be enumerated as follows:

- a. to provide a linguistically faithful and scientifically constructed image of spoken French, in both its unity and diversity, be it social, stylistic, or geographical;
- b. to enable researchers to test the hypotheses and phonological models, whether old or new, that are proposed for French, both in synchronic and in diachronic terms;
- c. to construct a representative database on the basis of a common and standardised methodology which allows secondary analyses to be undertaken in a variety of theoretical frameworks;
- d. to provide new reference data for applications in the domain of automatic processing of speech, French teaching and French linguistics.

Currently containing more than 900,000 words, the PFC is one of the largest audio-oral databases in the world. Over the course of the past ten years, 169 researchers have participated in the PFC programme (empirical enquiries, transcription, coding etc.).³ In 2010, thirty-three geographical regions of the French-speaking world accounting for seventy-six locations of enquiry were involved. Among the

2. In addition to France, PFC covers Belgium, Switzerland, Canada, Louisiana, the Maghreb, the Near and the Middle East, Africa, the French Caribbean Islands, the Indian Ocean, the Pacific. All information, data, protocols, the list of researchers and research teams involved, and the results of the programme can be accessed on the PFC website (www.projet-pfc.net).

3. The programme itself draws on the support of twelve different research teams and has benefited from various funding sources. More than ten doctoral projects have been either partly or entirely carried out within the framework of the programme, and five others are still in progress. More than 150 publications concerning the whole domain of French phonology have come out from the PFC programme.

seventy-six enquiries, thirty-three are now terminated, and are available in the PFC database. In addition, thirty-three are still being processed (transcription, coding, etc.) and ten are currently being completed. In all, the PFC base now boasts recordings of 489 French speakers representing a total of around 730 hours of digitally recorded and indexed speech currently available online. A total of forty-one hours are transcribed, aligned and coded for liaison and for schwa.

Phonetic, phonological, lexical and discursive data have been collected by means of a series of Labovian enquiries (see for more recent approaches Labov et al. 2006), in which each speaker is classified according to gender, age, and socio-economic status. For each speaker, different types of oral production have been recorded: word list reading, text reading, guided one-to-one conversation, and free/spontaneous conversation with peers.

With specific regard to liaison, 49,728 sites of potential realization have been coded according to the relevant segmental and contextual information. This comprehensive and theoretically neuter coding allows the researchers to have direct access to multiple sorting possibilities.

Located at the intersection of at least four contradictory historical dynamics, liaison is a very complex phenomenon (Durand et al. 2011). The historical dynamics leading to the fall of the final consonant since late antiquity Latin has fostered open syllabification. In contrast, generalized linkage, insofar as it fosters long phonological sequences, promotes the maintenance of these final consonants before an initial vowel, as a support for CV linkages. Furthermore, some of these consonants take on a role as markers for number and person, and thus tend to elude the dynamics of erasure: they tend to be conserved not in the coda of the final syllable of the word which they mark, but rather, as onsets of the opening syllable of the following word, thus undergoing resyllabification. Finally, spelling conventions demand that the etymological final consonant be consistently transcribed, whether it is realized in pronunciation or completely lost (Laks 2005, 2006, 2011; Clédât 1917).

Although generative and post-generative phonology (Schane 1965; Dell 1973; Encrevé 1988; Tranel 1995a, 1995b) has proposed that French liaison is a homogenous phenomenon that can be represented as a single process whereby a variable surface erasure rule applies to abstract underlying consonants, more recent work has shown that liaison is a multifactorial and multilevel phenomenon highly sensitive to frequency effects (de Jong 1994; Fougeron et al. 2001; Laks 2007; Durand et al. 2011).

In particular, the PFC database has allowed large-scale investigations concerning all major factors of sociolinguistic and geographical variability, including age, gender, French spoken as a first vs. second language, diatopic variation of northern

vs. southern French etc. (e.g., Pagliano & Laks 2005; Durand & Lyche 2008; Mallet 2008; Coquillon et al. 2010; Durand et al. 2011). Some of these factors have been shown to have an impact on the frequency of realization of liaison in facultative contexts, as in the case of the age factor (old speakers tend to produce more liaisons than younger speakers, Durand et al. 2011), while other factors have been proved not to influence the use of liaison substantially (as in the case of geographic variation and gender, Coquillon et al. 2010). Finally there are factors that still need to be thoroughly investigated; one such example is the educational level of the speakers (see Durand et al. 2011: 121). Influence of socio-economic status still deserve precise investigations based on large data bases; see for example the recent analyses of the influence of parental socio-economic status on children acquisition of liaison by Chevrot et al. (2011), and Hornsby's (2012) stratificational analysis of liaison errors and repairs.

In this paper, we propose a frequency analysis of French liaison and liaison environments focusing on the specific contexts of liaison realization attested in the PFC database. Liaison environments are the attested combinations of two consecutive words generating liaison at their juncture. The frequency analysis is aimed to uncover the distributional aspects of French liaison in its actual lexical instantiations, under the general usage-based hypothesis that liaison is more frequently realized in those word groups that have strong internal cohesion and high frequency of co-occurrence (e.g., Bybee 1998, 2001, 2005).

We will verify this distributional hypothesis over a very large amount of productions realized by different groups of French speakers varying in terms of age and educational level. The findings will show that our initial hypothesis is correct inasmuch as the distribution of liaison is similar to a power-law distribution in which a few types are ranked high for productivity and account for approximately one-half of the total observations. We will see that this generalization holds true for all types of linking consonants and all groups of speakers (as determined according to age and educational level), with some oscillations concerning different types of infrequent liaison patterns.

The paper is structured as follows. Section 2 presents our working definition of liaison environment and illustrates aims and procedures of the distributional analysis of French liaison in PFC. Section 3 presents the main results separately for the PFC corpus (§3.1), for some individual liaison consonants (§3.2) and for some subgroups of speakers as defined by age and educational level (§3.3). Section 4 presents the general discussion and concludes.

2. The distributional analysis of French liaison

This analysis deals with 16,805 enacted liaisons produced in either free or guided conversations.

Each enacted liaison corresponds to a liaison environment, comprising a left word ending in a consonant and a right word beginning with a vowel. Although in very rare cases the liaison consonant may be realized as unlinked (that is, with the consonant realized as the coda of the first word; Encrevé 1988), the present analysis deals with linked consonants only. The total repertory of liaisons therefore includes all liaison environments with their frequencies of occurrence. As our aim here is to analyze the productivity power of each type of environment by calculating the number of realizations produced, we concentrate here on enacted linking; we postpone the ratio analysis of realized vs. virtual liaison to a forthcoming paper on the social stratification of French liaison.

Figure 1 provides an example of a liaison environment. Recording ID 75cv11 corresponds to the performance <trois> *linked* <ans> [trwazā]. The lexical context for this particular occurrence enables us to construct the environment type <trois> *linked* <ans>. If the combination <trois> *linked* <ans> appears in other recordings of the PFC corpus, the frequency of occurrence (or token frequency) of this liaison environment will be > 1.

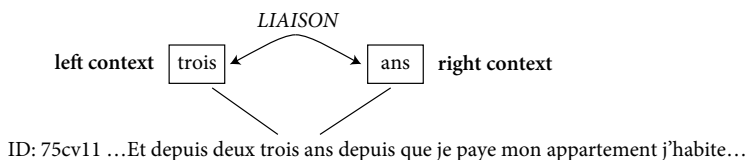


Figure 1. A liaison environment.

We ranked all liaisons' environments attested in the PFC corpus according to their token frequency. Next, the frequency cumulative percent of all environments was calculated to evaluate the productivity of each individual environment and of the subgroups of these environments.

3. Results

3.1 Distributional analysis of liaison types

The 16,805 liaison occurrences produced in free and guided conversations turned out to be organized in 3,105 environments (or “types”) of liaison. Each environment was defined by a given token frequency, ranging from 1,318 to 1. The data were plotted into a log-log graph (Figure 2). Log-log graphs are two-dimensional graphs of numerical data that use logarithmic scales on both the horizontal and vertical axes, and can be used to examine the tail of a distribution of data.

In statistics and probability theory the use of a log-log graph for plotting data distribution is a common practice because it allows for clear visualization even for data which is scarce in frequency.

In our analysis we plot the frequency of each liaison type along the y-axis and along the x-axis we report the rank of each type according to their frequency.

If the points in the plot tend to converge into a straight line for large numbers in the x-axis, then the researcher concludes that the distribution has a power-law tail (Jeong et al. 2000). Figure 2 displays the rank order of each liaison type by its number of occurrences in the corpus (y-axis).

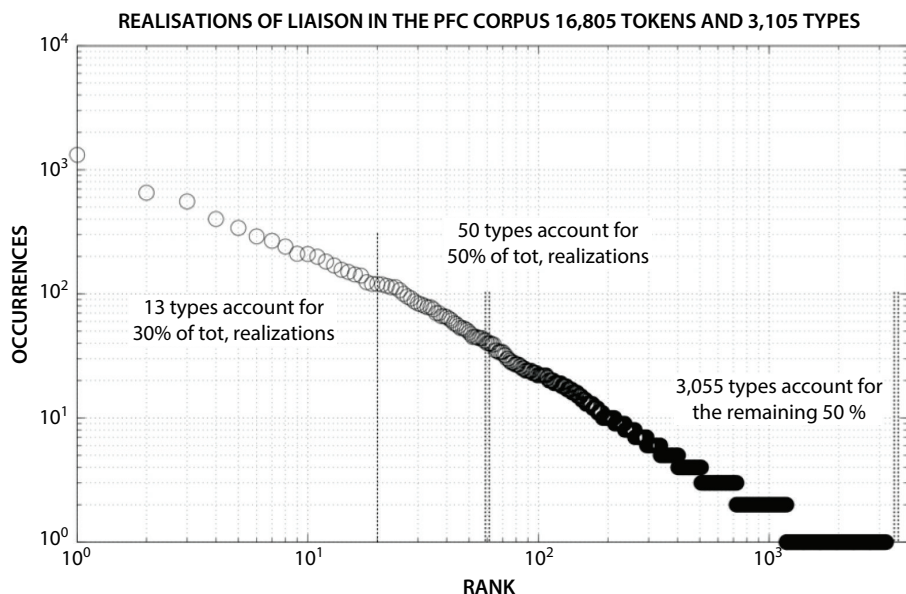


Figure 2. Log-log plot of liaison environments (or ‘types’) in the PFC corpus (rank order by number of occurrences).

We recognize in Figure 2 what is a typical distribution, first described by Zipf (1949) and later generalized by Mandelbrot as the Mandelbrot-Zipf distribution (Brillouin 1959). This typical distribution derives from the power-law and is close to the distribution pattern described at the end of the 19th century by Pareto. When the frequency with which an event occurs varies as a power of some attribute of that event (e.g. its size), the frequency is said to follow a power-law. In linguistics, one famous example of power-law functions is the Zipf's law in corpus analysis, according to which the frequency of a word item in a text is inversely proportional to its frequency rank (i.e., the second most frequent word item occurring half as often the most frequent item; Zipf 1949).

In the case of Mandelbrot-Zipf distributions, a clear distinction emerges between two zones of the curve (Wimmer & Altman 1999): the peak (or head) zone and the dispersal (or tail) zone. In the peak zone, a very small number of highly productive events are concentrated. The peak zone thus differs very neatly from the dispersal zone, an asymptotically zero zone where a large number of infrequent events with a marginal impact on the process are distributed. The latter condition is known as the phenomenon of the 'long tail'. The 'body' of the curve is represented by a gradual shift from the peak to the dispersal distribution.

The frequency analysis confirmed that the distribution of the lexical environments defining French liaison follows a power-law distribution to a significant extent. We calculated a goodness-of-fit score using the method of maximum likelihood in order to estimate the scaling exponent and the lower bound value of the distribution according to Clauset et al. (2009). The maximum likelihood estimators converge on the correct value of the scaling exponent, with probability 1 for both the discrete and the continuous power-laws (Clauset et al. 2009). Figure 3 shows the cumulative distribution function of the frequency of the lexical components involved in the French liaison and its fitted power-laws function on the basis of the maximum likelihood estimators. The fitted power-laws function was subsequently tested by calculating the goodness-of-fit with the actual liaison data. We used the Kolmogorov-Smirnov test (Clauset et al. 2009: 14) to calculate the p value of the goodness-of-fit function between the power-laws distribution and the liaison data as reported in Figure 3. The p value should be greater than 0.1 to allow for a plausible hypothesis on the data, otherwise the hypothesis has to be rejected. The obtained value $p=0.111$ indicated that the distribution of the lexical environments defining French liaison follows a power-laws distribution to a statistically significant extent.

Thus, according to the results of this global inspection of PFC data an extremely small number of types represent quantitatively the core of the process of French liaison, whereas the entire set of the remaining types accounts for no more than half of the actual occurrences.

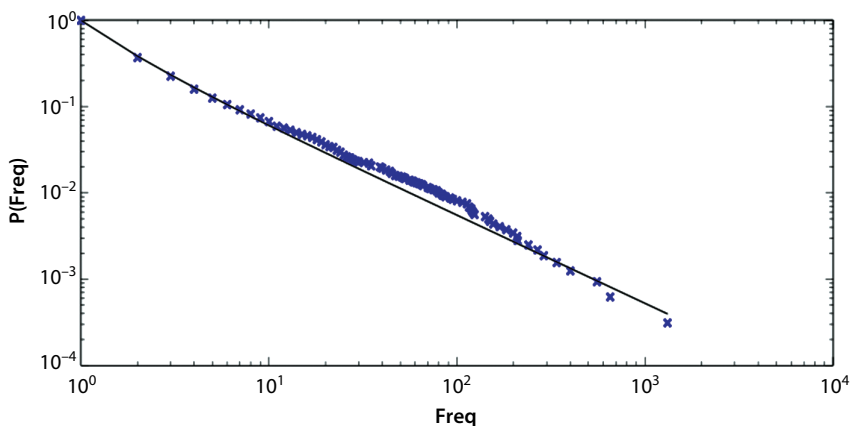


Figure 3. Cumulative distribution function for the lexical environments defining French liaison in the PFC corpus and fitted power-law distribution function obtained through maximum likelihood estimators.

More specifically, the 13 most frequent types alone represent the 30% of the 16,805 occurrences of enacted liaison, while as few as 50 types account for half of the total number of liaisons. The distribution of those high-frequency types is very sparse inasmuch as there is a relatively sharp decline of the number of occurrences in moving across the 13 most frequent types. The tail of the distribution is equally remarkable, but for the opposite reason: no less than 3,055 types must be called upon to cover about the same number of realized liaisons, i.e. the remaining 50% of the realizations. The distribution of those 3,055 types in the curve is crowded, as opposed to the sparseness of the head, since many types show the same low frequency values and the farther right we move along the curve, the more types with the same frequency value are found. For this reason the tail of the curve is much milder than the head and it ends up flat.

In other words, liaison types occupying the tail of the curve have a very low probability of occurrence in the corpus but since they are very numerous, they are equally essential for the global picture.

Table 1 illustrates the nature of the two zones of the curve with some examples. Few types at the top of the frequency ranking yield a relatively high cumulative percent. On the contrary the bottom zone of the ranking (corresponding to the long-tail of the curve) is occupied by a very large number of very rare liaison types: almost 1,800 liaison types have a token frequency equal to 1.

These data provide interesting challenges to usage-based and exemplarist models of phonological processing, in which lexicon and grammar are integrated and constrained by the same organizational principles (e.g., Bybee 1998;

Table 1. Number of occurrences (frequency), percent and cumulative percent for the 10 highest ranked the 13 lowest ranked liaison types in the PFC corpus.

Realizations of liaison in the PFC corpus 16,805 tokens and 3,105 types				
Rank	Liaison	Frequency	Percent	Cumulative percent
1	on_L_a	1318	7.84	7.84
2	on_L_est	653	3.89	11.73
3	ils_L_ont	555	3.30	15.03
4	en_L_a	401	2.39	17.42
5	on_L_avait	339	2.02	19.43
6	on_L_était	290	1.73	21.16
7	quand_L_on	268	1.59	22.76
8	dans_L_un	240	1.43	24.18
9	deux_L_ans	210	1.25	25.43
10	dans_L_une	209	1.24	26.68
...
3093	vous_L_apporte	1	0.01	99.93
3094	vous_L_apprenait	1	0.01	99.93
3095	vous_L_apprend	1	0.01	99.94
3096	vous_L_apprendre	1	0.01	99.95
3097	vous_L_approvisionner	1	0.01	99.95
3098	vous_L_arrêter	1	0.01	99.96
3099	vous_L_arrêtez	1	0.01	99.96
3100	vous_L_arriverez	1	0.01	99.97
3101	vous_L_arriviez	1	0.01	99.98
3102	vous_L_assurer	1	0.01	99.98
3103	vous_L_attendez	1	0.01	99.99
3104	vous_L_auriez	1	0.01	99.99
3105	vous_L_autres	1	0.01	100.00

Langacker 1987). In particular, the phenomenon of the ‘long tail’ asks for a refinement of the view that part of the liaison process must be inscribed as a repository of ‘constructions’ in the mental lexicon (Bybee 2005; Bybee and McClelland 2005). According to the latter view, ‘irregular’ liaison is stored in the lexical repository of constructions as a list of types, and “a certain level of token frequency is necessary to maintain these irregularities” (Bybee 2005: 24). This analysis fits well with the head of the curve: tracing the 13 lexical environments (see Figure 2) which account for 30% of the overall data back to a set of derivational rules cannot in any case be shorter than providing the exhaustive list which accounts for them. However, the tail of the distribution cannot be dealt with by the same cognitive mechanism of full-form storage: most of the types that we find here have a global

frequency that is only slightly higher than or equal to 1. Therefore, memorization alone cannot account for the whole story. And, contrary to Bybee's (2005) prediction that low frequency constructions tend to be lost, what we found here is that a very long list of very infrequent lexical environments account for 40–50% of the occurrences of liaison in the corpus.

We suggest therefore that a productive process of generalization has to be postulated, in order to account for the existence of such a long list of very infrequent liaison constructions. This is a potentially open list, inasmuch as some relevant linguistic feature (either morphosyntactic, or phonological, or both, according to the individual cases) could in principle trigger the process of generalization even further.

3.2 Distributional analysis of liaison consonants

As a further step in the analysis, we wanted to evaluate whether the obtained distribution is consistently represented in some relevant subgroups of data defined by 'system-internal' and 'system-external' factors. As an internal factor we considered the phonological nature of the linking consonant, limited to the three most frequent consonant types in the corpus, which are /n/, /z/ and /t/, respectively. As external factors we included the two sociolinguistic factors of age and educational level, as encoded in the PFC corpus. The former is known to affect slightly the production of liaison inasmuch as older speakers are generally said to produce more liaisons in facultative contexts than younger speakers (e.g., Malécot 1975; Ashby 1981; Booij & de Jong 1987; Ranson 2008; Durand et al. 2011). The latter still is a poorly investigated factor (Durand et al. 2011: 121) and the PFC database appears to have strong potential for future research. By introducing phonological and sociolinguistic variables into the picture, whose influence on liaison production is either well-known or still a matter of debate, we wanted to verify whether the power-laws statistical distribution of enacted liaisons varies according to different subgroups of data or whether, on the contrary, it resists any manipulation of the corpus.

Figure 4 illustrates the data split by consonant type (with /n/, /z/ and /t/ as the relevant consonants) compared to the global distribution of the liaison data. The top of the figure is occupied by a reduced version of Figure 2, reproduced here to facilitate internal comparisons. The three bottom diagrams refer to the liaison environments realized by /n/, /t/ and /z/, respectively. The sibilant fricative is the most frequent liaison consonant in the corpus with 7,840 occurrences (corresponding to the 47% of the corpus). The nasal is also extremely present in the corpus (6,449 occurrences, corresponding to the 39% of the corpus), while the alveolar stop is present in a smaller proportion (2,342 occurrences, almost

the 14% of the corpus). The lexical environments represented in each of the three bottom diagrams are mutually exclusive, inasmuch as each lexical environment is (by definition) realized by one specific liaison consonant.

One can see from the figure that, although /n/ has more data points in the diagram compared to /t/ and /z/ is even more frequent than /n/, the global distribution is similar for the three consonants. The head of the curve is occupied by a relatively limited number of high-frequency types, with discrete frequency values. The 30 most frequent lexical environments are shown in Table 2 according to consonant type. Five out of the six most frequent types contain /n/, and only one contains /z/. There is a strong lexical bias among these high frequency types, since four of the five /n/ types are realized by the indefinite 3rd person pronoun *on*. On the other hand, 12 of the 25 most frequent environments contain a /z/. This indicates that, although /z/ is more frequent than /n/ as a liaison consonant overall, the liaison environments located very high in the frequency curve are specified by /n/, not /z/. We must therefore exclude those few high-frequency /n/ environments, in order to get a more balanced picture across consonant types.

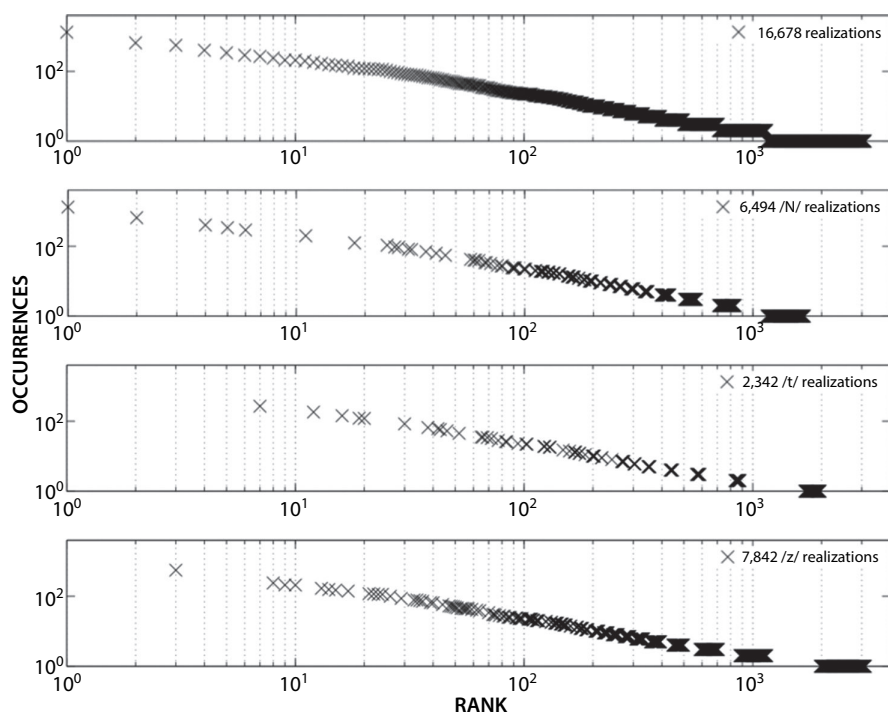


Figure 4 (a–d). Log-log plot of liaison environments (or ‘types’) in the PFC corpus (rank order by number of occurrences) (a); split by type of the liaison consonant (b–d).

On the other hand, as long as frequency decreases, the number of lexical environments per frequency value increases, and the curve ends up in a flat distribution where many liaison types feature a frequency value equal to 1. The latter group of lexical environments with frequency = 1 is obviously more numerous

Table 2. Number of occurrences (frequency), percent and cumulative percent for the 30 highest ranked liaison types in the PFC corpus, separately for the three consonants /n/, /t/ and /z/.

Realizations of liaison specified by consonants /n/, /t/ and /z/						
Rank	Liaison	Frequency	Cumulative percent	/n/ consonant	/t/ consonant	/z/ consonant
1	on_L_a	1318	7.8	1318	0	0
2	on_L_est	653	11.7	653	0	0
3	ils_L_ont	555	15.0	0	0	555
4	en_L_a	401	17.4	401	0	0
5	on_L_avait	339	19.4	339	0	0
6	on_L_était	290	21.2	290	0	0
7	quand_L_on	268	22.8	0	268	0
8	dans_L_un	240	24.2	0	0	240
9	deux_L_ans	210	25.4	0	0	210
10	dans_L_une	209	26.7	0	0	209
11	on_L_a	199	27.9	199	0	0
12	est_L_un	183	28.9	0	183	0
13	les_L_enfants	169	30.0	0	0	169
14	vous_L_avez	156	30.9	0	0	156
15	trois_L_ans	150	31.8	0	0	150
16	tout_L_à	143	32.6	0	143	0
17	les_L_autres	141	33.5	0	0	141
18	n_L_ai	124	34.2	124	0	0
19	quand_L_il	122	34.9	0	122	0
20	est_L_une	120	35.6	0	120	0
21	ils_L_avaient	119	36.4	0	0	119
22	ils_L_étaient	116	37.0	0	0	116
23	nous_L_a	113	37.7	0	0	113
24	vous_L_etes	113	38.4	0	0	113
25	un_L_an	107	39.0	107	0	0
26	nous_L_avons	100	39.6	0	0	100
27	on_L_en	95	40.2	95	0	0
28	on_L_allait	92	40.7	92	0	0
29	des_L_études	87	41.2	0	0	87
30	quand_L_ils	84	41.7	0	84	0

for /z/ and /n/ than for /t/, as was the group of lexical environments occupying the head positions. Table 2 shows that the alveolar stop is the least represented consonant if one considers the 30 most frequent environments; however, the same holds true even if one considers the tail of the curve. In sum, the body and the tail of the curve do not reveal any significant difference in the distribution of liaison environments across consonant categories.

From this analysis we can conclude that the tendency towards the power-laws distribution that we have observed over the whole repository of forms remains stable overall even if we consider different phonological contexts of enacted liaisons separately. There are however some idiosyncrasies in the head zone related to the fact that the nasal is the liaison consonant occupying the five highest positions in the frequency curve, and in particular, the lexical environments including *on + a*, *est*, *avait*, *était* stand out because of their enormous number of occurrences in the corpus. This distributional bias accounts for the particular behaviour of the nasal consonant in the head zone, compared to /z/ and /t/.

3.3 Distributional analysis of liaison types according to age and educational level

By considering the productions of different groups of speakers we wanted to verify whether any of the liaison environments is specific to any group of speakers in such a way that the general picture of the power-laws distribution disappears when liaison is analysed with respect to particular subsets of linguistic use. These subsets are defined here in terms of some basic, broadly conceived sociolinguistic factors, that could in principle influence the linguistic behavior of the speakers with respect to the production of liaison (see above, §1, for some background on sociolinguistic variability in the production of French liaison). As specified above, we took age as a typical sociolinguistic factor generally considered to influence non-obligatory liaison production. On the contrary, educational level was included in the analysis because it represents a factor which is still poorly investigated. As for age, we divided the PFC speakers into two groups: those aged 50 or below, and those older than 50. We are conscious of the fact that this bipartition is not as fine-grained as a sociophonological inquiry would require, but there were limitations in the distribution of the PFC speakers across the different age classes that forced us to adopt such a neat bipartition. As for educational level, we created three groups of speakers: those with up to 14 years of scholastic education ('low' educational level henceforth), those with up to 18 years ('intermediate' educational level) and those with up to 24 years ('high' educational level).

Figure 5 illustrates the data split by age groups compared to the global distribution of the liaison data. The two bottom diagrams refer to the liaison environments realized by the ≤ 50 and > 50 year-old speakers, respectively, while the top of the figure refers to the total amount of liaison environments resulting from the sum of the two age subgroups. Differently from the analysis by consonant, the lexical environments represented in each of the two bottom diagrams are not mutually exclusive by definition, since each lexical environment can be (and often is) repeated by the speakers of both groups.

The diagrams showed that the global distribution of the data across the curve was approximately the same for the two age groups and no significant idiosyncrasies could be found: as repeatedly observed, the head of the curve is occupied by the same small number of high-frequency types, with discrete frequency values. Table 3 shows that the identity and distribution of those high-frequency lexical types tended to be consistent across the two age groups. On the contrary, as far as frequency decreases, the number of lexical environments per frequency value increases. The curve ends up in a flat distribution where many liaison types feature a frequency value equal to 1. However, the tail zone shows that there are differences according to the age of the speakers and some types featuring a frequency value equal to or lower than 10 are evenly distributed across the two classes. This is particularly evident in looking at Figure 6, which presents a magnification of

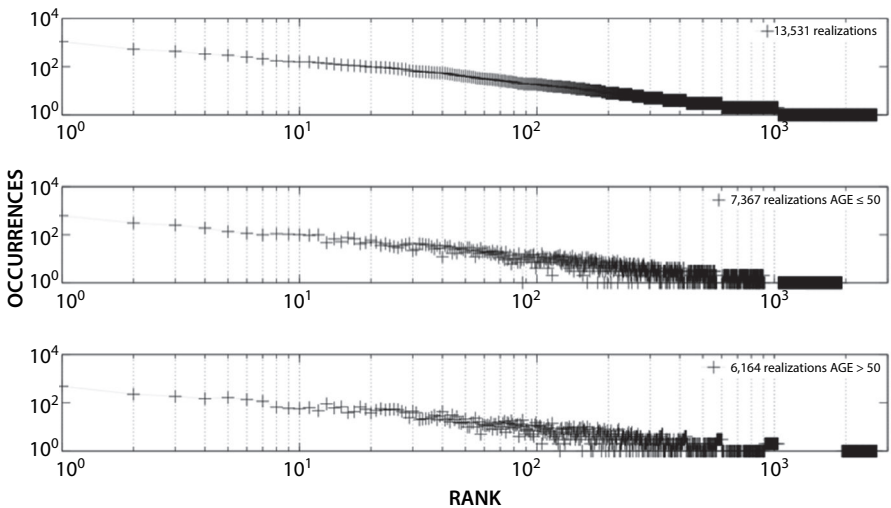


Figure 5 (a-c). Log-log plot of liaison environments (or ‘types’) in the PFC corpus (rank order by number of occurrences) (a); split by AGE ≤ 50 and AGE > 50 (b-c).

Table 3. Number of occurrences (frequency), percent and cumulative percent for the 25 highest ranked liaison types in the PFC corpus, separately for AGE ≤ 50 and AGE > 50.

Realizations of liaison specified by age					
AGE ≤ 50 and AGE > 50					
Rank	Liaison	Frequency	Cumulative percent	AGE ≤ 50	AGE > 50
1	on_L_a	1098	8.1	617	481
2	on_L_est	530	12.0	306	224
3	ils_L_ont	434	15.2	250	184
4	en_L_a	338	17.7	190	148
5	on_L_avait	299	19.9	135	164
6	on_L_était	251	21.8	114	137
7	quand_L_on	214	23.4	99	115
8	dans_L_un	175	24.7	109	66
9	dans_L_une	164	25.9	105	59
10	deux_L_ans	158	27.1	101	57
11	est_L_en	158	28.2	95	63
12	les_L_enfants	146	29.3	99	47
13	est_L_un	137	30.3	47	90
14	vous_L_avez	128	31.3	67	61
15	tout_L_à	119	32.1	51	68
16	les_L_autres	114	33.0	75	39
17	trois_L_ans	111	33.8	58	43
18	quand_L_il	109	34.6	42	67
19	vous_L_etes	101	35.4	47	54
20	ils_L_étaient	97	36.1	59	38
21	n_L_ai	97	36.8	49	48
22	est_L_une	93	37.5	36	57
23	nous_L_a	93	38.2	43	50
24	nous_L_avons	88	38.8	35	53
25	ils_L_avaient	86	39.5	32	54

the tail of the curve. Some low-frequency realizations are typically produced by younger speakers, some others are typically produced by older ones. The token frequency value of these evenly distributed lexical environments is low, but they are numerous. The resulting picture is visibly differentiated across age groups for what concerns the selection of low-frequency lexical environments featuring liaison.

The analysis of the liaison distribution according to the speakers' educational level is conducted in the same way as the analysis by age. Figure 7 illustrates the data split by educational level (three bottom diagrams) compared to the global distribution of the liaison data resulting from the summation of the three subgroups.

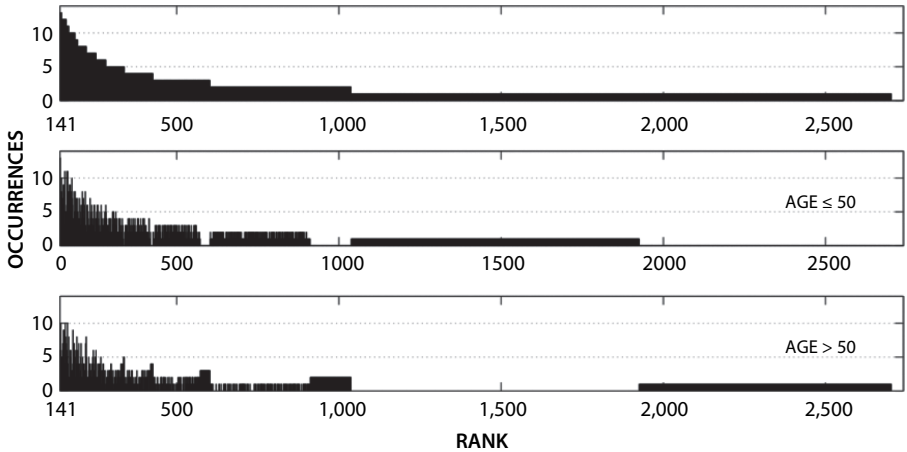


Figure 6 (a–c). Magnification of the tail for AGE ≤ 50 and AGE > 50 .

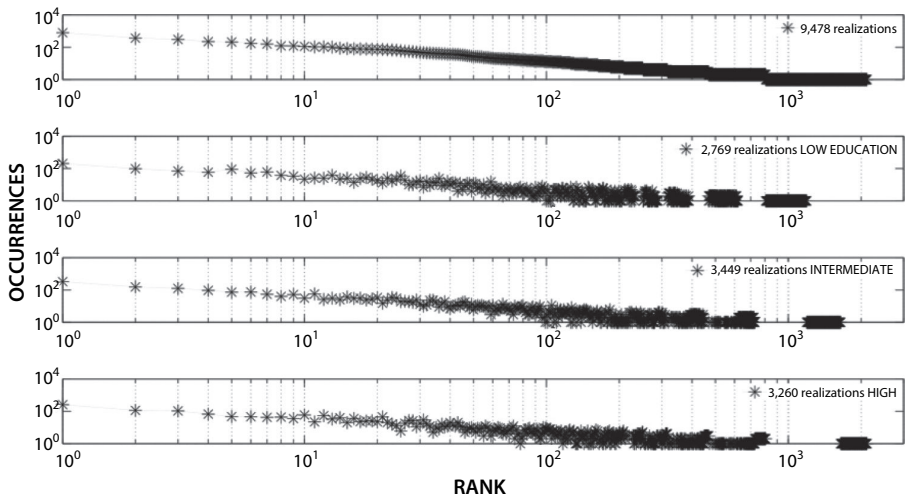


Figure 7 (a–d). Log-log plot of liaison environments (or ‘types’) in the PFC corpus (rank order by number of occurrences) (a); split by education level (b–d).

As in the case of the analysis by age groups, the lexical environments represented in each of the three bottom diagrams are not mutually exclusive by definition. Table 4 illustrates the twenty-five most frequent lexical environments occupying the head of the curve. Figure 8 is a magnified representation of the tail of the curve.

As in the case of age grouping discussed above, the diagrams relative to educational level showed that the global distribution of the data across the curve was approximately the same for the three groups, particularly as far as the head was concerned, while the tail showed that some of the types featuring a frequency value equal to or lower than 10 were evenly distributed across the three groups. These particular environments differentiated the production of the three groups and indicated that variation in the low-frequency zone may arise according to the social characterization of the speakers.

Table 4. Number of occurrences (frequency), percent and cumulative percent for the 25 highest ranked liaison types in the PFC corpus, separately for the low, intermediate and high education levels.

Realizations of liaison specified by education level LOW, INTERMEDIATE and HIGH LEVEL						
Rank	Liaison	Frequency	Cumulative percent	LOW	INTER- MEDIATE	HIGH
1	on_L_a	786	8.3	204	326	256
2	on_L_est	367	12.2	97	155	115
3	ils_L_ont	301	15.3	70	125	106
4	en_L_a	220	2.3	59	94	67
5	on_L_avait	210	19.9	90	73	47
6	on_L_était	174	1.8	53	75	46
7	quand_L_on	157	23.4	61	54	42
8	dans_L_un	122	24.7	38	40	44
9	est_L_a	121	25.9	34	52	35
10	dans_L_une	112	27.1	22	31	59
11	est_L_un	105	28.2	26	57	22
12	vous_L_avez	105	29.3	25	26	54
13	deux_L_ans	102	30.4	38	31	33
14	les_L_enfants	87	31.3	22	25	40
15	quand_L_il	82	32.2	25	37	20
16	les_L_autres	80	33.0	14	30	36
17	ils_L_étaient	78	33.9	24	31	23
18	tout_L_à	77	34.7	28	24	25
19	trois_L_ans	75	35.5	22	27	26
20	est_L_une	74	36.2	18	34	22
21	vous_L_etes	73	37.0	13	15	45
22	n_L_ai	69	37.7	25	24	20
23	nous_L_a	67	38.4	17	34	16
24	ils_L_avaient	66	39.1	22	31	13
25	on_L_allait	60	39.8	35	19	6

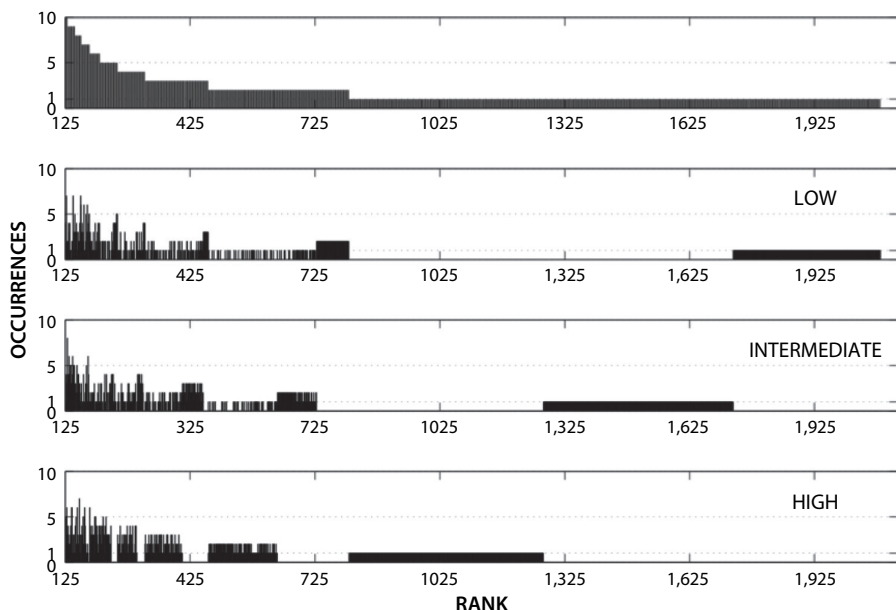


Figure 8. Magnification of the tail for the low, intermediate and high education levels.

Taken together, the analyses by age and educational level clearly indicate that French liaison features a statistical tendency toward a power-laws distribution independently of the social characteristics of the speakers. This finding substantially agrees with the results of the analysis by consonant class (see above, §3.2) inasmuch as the power-laws distribution of enacted liaisons turns out to resist any statistical manipulation of the investigated corpus. On the other hand, splitting the data by age or educational level has revealed that different groups of speakers produce different examples of low frequency liaison environments. Our analysis does not reveal anything about whether this variability is correlated with particular social factors. Studies specifically devoted to illustrating the fine-grained patterns of variation according to these and other social factors will be more revealing in this respect. However, our analysis does reveal that if we want to analyse the nature and extent of the variability associated with the production of liaison, we should preferably look at the tail of the distribution, rather than at its head (or its body). The variegated sample of low- and very-low-frequency items is the most likely repository of lexical environments differentially selected by different groups of speakers.

The size of the corpus over which the analysis is conducted is therefore crucial in this respect. Our PFC sample has allowed us to uncover the different behaviour of the head vs. the tail of the curve because it includes such a large number of

individual realizations, distributed over a socially and geographically stratified collection of speakers. The analysis of enacted liaison in the PFC sample thus suggests that corpus analysis may reveal the existence of subtle variations even for those phenomena generally considered as ‘monoliths’ in a linguistic repertoire and with very few internal sociophonological variation.

4. General discussion

At the conclusion of this quantitative analysis of liaison in the PFC corpus, the image which emerges is one which differs significantly from that put forward by the generative or post-generative phonologies which postulate the presence in all cases of a latent consonant in an abstract representation. We have shown that if we decide to analyze not the absence of liaison (as a consequence of the erasure process) but rather the process that positively produces liaison, we find that a relatively small number of <word> *linked* <word> constructions may account for up to the 50% of total productions. In agreement with Bybee’s suggestions about the statistical organization of the mental lexicon (Bybee 2005, 2006; Bybee and McClelland 2005) and the theories of Usage Grammars (Barlow and Kemmer 2000; Langacker 2000), we believe that these constructions should be regarded as ‘frozen’, i.e., stored as such in the mental lexicon (see also Chevrot et al. 2011 for an acquisitional point of view). However, we have seen that mnesic storage of recurrent <word> *linked* <word> constructions, probably to be considered as single words with lexicalized liaison, is not in itself sufficient to explain liaison and its complexity. We must also postulate the existence of a more marginal process, allowing the generation of a very large number of those rare liaisons that represent the remaining 50% of all enacted liaisons. Mnesic storage corresponds to a generalized linkage dynamic that is typical of cursus languages (Pulgram 1970). Such a process erases the boundaries between words and confers upon the prosodic group a central role in oral production (Grammont 1914). By contrast, generalization of rare liaisons is a tendency driven by the orthographical demarcation between words. This process, in contrast to mnesic storage, reaffirms the orthographical integrity of words.

It is also important to remark that traditional lexicalist analyses of French liaison have been based on the analysis of “constructions”, i.e., “groupes de mots qui ont une ‘forte cohésion syntaxique’” (Bybee 2005: 24). These are therefore grammatical constructions, instantiated by specific lexical sequences with their own degree of syntactic cohesion (usually strong to very strong) and their own frequency of co-occurrence (varying from low to high). Syntactic cohesion and frequency of co-occurrence are the epiphenomena of lexical properties of

the words included in the constructions such as their “relations sémantiques, fonctionnelles ou statistiques particulières lorsqu’ils [= the words] sont utilisés ensemble” (Bybee 2005: 26). One very typical example of liaison construction is [DETERMINANT + NOUN], as e.g. in <les> linked <enfants> (this particular instantiation being a relatively high-frequency lexical environments, according to our own analysis, as it occupies position 13th in Table 1). However, syntactic cohesion and frequency of co-occurrence sometimes conflict. One such case is reported in Bybee (2005: 35): the construction [VERB + DETERMINANT] in the specific case of <est> linked <un> (followed by a noun) generates much more liaisons than any other contexts in which <est> co-occurs with other vowel-initial words, including those with a higher degree of syntactic cohesion (such as [AUXILIARY + PAST PARTICIPLE] <est> linked <arrivé>, for instance). Bybee is therefore forced to conclude that “[c]ette tendance suggère fortement qu’il existe une construction dans laquelle *est* [t] *un* est un constituant qui précède un nom” (Bybee 2005: 26).

Our analysis by making an appeal to the notion of lexical environment and completely disregarding the convention of referring each lexical environment to a general morpho-syntactic label, overcomes this *impasse* and proposes a statistical approach in which liaison environments are exclusively specified by the frequency of their usage in the corpus. On this basis, our analysis emphasizes the statistical distribution of actual usages in a very large repository of forms, with no superimposed predictions about how and to what extent syntactic cohesion produces secondary assemblages within the classes of exemplars (Foulkes 2006).

5. Conclusions

Starting from the acknowledgment of corpus phonology as a truly sociophonological practice, we have proposed a frequency analysis of French liaison focusing on the specific contexts of liaison realization attested in the PFC database. All attested combinations of two consecutive words generating liaison at their juncture were analyzed to uncover the distributional aspects of French liaison in its actual lexical instantiations.

The results of the analysis, indicating the existence of a power-laws distribution in the production of liaison, suggested that there is a complex function regulating the number of liaison environments and their token frequency in a corpus. The phonological component, intended as the nature of the liaison consonant, does not add much information to the analysis of such distribution, since very few inconsistencies are found across different consonants. External factors, such as the age and the educational level of the speakers, neither provide substantially

different pictures of the general distribution function; they do, however, provide some suggestion about the need for a closer inspection of corpora concerning in particular low- or very-low-frequency lexical uses. We therefore take this point as one of the fundamental lesson of our corpus study, allegedly bearing important consequences for current sociophonological practice: there are phenomena in the phonology of languages, for which a “core” of frequent lexical uses may be substantially untouched by sociolinguistic variation, while a “periphery” of infrequent uses appears to show significant aspects of style- or speaker-dependent variation. The methodological corollary is the importance of basing any variationist analysis on very large data sample, such as those provided by contemporary, well-reasoned linguistic corpora.

Among the future perspectives of this work is the analysis of the distributional differences (or similarities) between obligatory and facultative liaison. Such differentiation will also allow us to approach, within our distributional hypothesis, the issue of realized vs. potential but non realized liaison. According to the methodology inaugurated in the present study, the two types of realized liaison will be analyzed over the whole dataset of attested liaisons first, and subsequently as a function of basic sociolinguistic variables. It has to be verified, as a matter of fact, whether sociolinguistic variation affects liaison production differently for the two different types of liaison (i.e., obligatory vs. facultative); recent studies do seem to suggest this (e.g., Chevrot et al. 2011 for children’s productions; Hornby 2012), and a distributional investigation across the PFC corpus is likely to uncover important aspects of stratification in language use.

References

- Ashby, William J. 1981. “French liaison as a sociolinguistic phenomenon”. *Linguistic Symposium on Romance Languages* ed. by William W. Cressey & Donna J. Napoli, 46–57. Washington: Georgetown University Press.
- Barlow, Michael & Suzanne Kemmer, eds. 2000. *Usage based models of language*. Stanford: CSLI.
- Boersma, Paul & Silke Hamann, eds. 2009 *Phonology in perception*. Berlin: Mouton de Gruyter.
- Booij, Geert & Daan de Jong. 1987. “The domain of liaison: theories and data”. *Linguistics* 25. 1005–1025. DOI: 10.1515/ling.1987.25.5.1005
- Brillouin, Léon. 1959. *La science et la théorie de l’information*. Paris: Masson.
- Bybee, Joan L. 1998. “A functionalist approach to grammar and its evolution”. *Evolution of Communication* 2. 249–278. DOI: 10.1075/eoc.2.2.06byb
- Bybee, Joan L. 2001. “Frequency effects on French liaison”. *Frequency and the emergence of linguistic structure* ed. by Joan L. Bybee & Paul Hopper, 337–359. Amsterdam & Philadelphia: John Benjamins.
- Bybee, Joan L. 2005. “La liaison: effets de fréquence et constructions”. *Langages* 158. 24–37. DOI: 10.3917/lang.158.0024

- Bybee, Joan L. 2006. *Frequency of use and the organization of language*. Oxford: Oxford University Press.
- Bybee, Joan L. & James L. McClellan. 2005. "Alternatives to the combinatorial paradigm of linguistic theory based on domain general principles of human cognition". *The Linguistic Review* 22/2–4. 381–410. DOI: 10.1515/tlir.2005.22.2-4.381
- Chevrot, Jean-Pierre, Aurélie Nardy & Stéphanie Barbu. 2011. "Development dynamics of SES-related differences in children's production of obligatory and variable phonological alternations". *Language Sciences* 33. 180–191. DOI: 10.1016/j.langsci.2010.08.007
- Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge: MIT Press.
- Clédat, Léon. 1917. *Manuel de phonétique et de morphologie historique du français*. Paris: Hachette.
- Clauset A., C. R. Shalizi, & M. E. J. Newman. 2009. "Power-law distributions in empirical data", *SIAM Review* 51/4. 661–703. DOI: 10.1137/070710111
- Coquillon, Annelise, Jacques Durand, Chantal Lyche & Julien Eychenne. 2010. "French Liaison: From global results to local varieties". *Colloque International PhonLex*, Toulouse, Université de Toulouse II Le Mirail, 8th–10th September 2010.
- de Jong, Daan. 1994. "La sociophonologie de la liaison orléanaise". *French Generative Phonology: Retrospective and Perspectives* ed. by Chantal Lyche, 95–131. Salford: ESRI.
- Dell, François. 1973. *Les règles et les sons : introduction à la phonologie générative*. Paris: Hermann.
- Durand, Jacques, Bernard Laks & Chantal Lyche. 2002. "La phonologie du français contemporain: usages, variétés et structures". *Romanistische Korpuslinguistik: Korpora und Gesprochene Sprache / Romance Corpus Linguistics: Corpora and Spoken Language* ed. by Claus D. Pusch & Wolfgang Raible. 93–106.
- Durand, Jacques, Bernard Laks & Chantal Lyche. 2005. "Un corpus numérisé pour la phonologie du français". *La linguistique de corpus* ed. by Geoffrey Williams, 205–217. Rennes: Presses Universitaires de Rennes.
- Durand, Jacques, Bernard Laks, Basilio Calderone & Atanas Tchobanov. 2011. "Que savons-nous de la liaison aujourd'hui?". *Langue Française* 169. 103–135. DOI: 10.3917/lf.169.0103
- Durand, Jacques & Chantal Lyche. 2008. "French liaison in the light of corpus data". *Journal of French Language Studies* 18/1. 33–66. DOI: 10.1017/S0959269507003158
- Encrevé, Pierre. 1988. *La liaison avec et sans enchaînement: phonologie tridimensionnelle et usages du français*. Paris: Le Seuil.
- Foulkes, Paul. 2006. "Phonological variation: a global perspective". *Handbook of English Linguistics* ed. by Bas Aarts & April M. S. McMahon, 625–669. Oxford: Blackwell. DOI: 10.1002/9780470753002.ch27
- Foulkes, Paul & Gerard J. Docherty. 2006. "The social life of phonetics and phonology". *Journal of Phonetics* 34/4. 409–438. DOI: 10.1016/j.wocn.2005.08.002
- Fougeron, Cécile, Jean-Philippe Goldman & Ulrich H. Frauenfelder. 2001. "Liaison and schwa deletion in French: an effect of lexical frequency and competition?" *EUROSPEECH 2001 – 7th European Conference on Speech Communication and Technology (Aalborg, September 3–7, 2001)* ed. by Paul Dalsgaard, Boslashge Lindberg, Henrik Benner, Zheng-Hua Tan, 639–642.
- Goldinger, Stephen D. 1998. "Echoes of echoes? an episodic theory of lexical access". *Psychological Review* 105. 251–279. DOI: 10.1037/0033-295X.105.2.251
- Grammont, Maurice. 1914. *Traité pratique de prononciation française*. Paris: Delagrave.

- Hornsby, David. 2012. "Getting It Wrong: Liaison, Pataquès, and Repair in Contemporary French". *On Linguistic Change in French: Socio-Historical Approaches. Festschrift for R. Anthony Lodge* ed. by Tim Pooley & Dominique Lagorgette, 69–84. Chambéry: Editions de l'Université de Savoie.
- Johnson, Keith. 2006. "Resonance in an exemplar-based lexicon: The emergence of social identity and phonology". *Journal of Phonetics* 34. 485–499. DOI: 10.1016/j.wocn.2005.08.004
- Labov, William. 2004. "Quantitative analysis of linguistic variation". *Sociolinguistics. An international handbook of the science of language and society*, Vol. 1, ed. by Ulrich Ammon, Norbert Dittmar, Klaus J. Mattheier & Peter Trudgill, 6–22. Berlin: de Gruyter.
- Labov, William, Sharon Ash & Charles Boberg. 2006. *The Atlas of North American English*. Berlin: Mouton De Gruyter.
- Laks, Bernard. 2005. "La liaison et l'illusion". *Langages* 158. 101–126. DOI: 10.3917/lang.158.0101
- Laks, Bernard. 2006. "Phonologie et construction syntaxique". *Linx* 53. 155–171.
- Laks, Bernard. 2007. "Les hommes politiques français et la liaison (1908–1999)". *Modéliser le changement : les voies du français* ed. by Luc Baronian & France Martineau. Montréal, Presses de l'Université de Montréal.
- Laks, Bernard. 2008. "Pour une phonologie de corpus". *Journal of French Language Studies* 18/1. 3–32. DOI: 10.1017/S0959269507003146
- Laks, Bernard. 2011. "La phonologie du français et les corpus". *Langue Française* 169. 3–9. DOI: 10.3917/lf.169.0003
- Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar I: Theoretical Prerequisite*. Stanford: Stanford University Press.
- Langacker, Ronald W. 2000. "A dynamic usage-based model". *Usage based models of language* ed. by Michael Barlow & Suzanne Kemmer, 1–65. Stanford: CSLI.
- Malécot, André. 1975. "French liaison as a function of grammatical, phonetic and paralinguistic variables". *Phonetica* 31. 161–179. DOI: 10.1159/000259697
- Mallet, Géraldine. 2008. *La liaison en français, descriptions et analyses dans le corpus PFC*. Thèse de doctorat, Université de Paris Ouest.
- Pagliano, Claudine & Bernard Laks. 2005. "Problématique de la liaison dans l'analyse d'un corpus de français oral actuel". *Colloque Français fondamental, corpus oraux, contenus d'enseignement. 50 ans de travaux et d'enjeux*, Ecole Normale Supérieure Lettres et Sciences Humaines, Lyon, 8–10 décembre 2005.
- Pulgram, Ernst. 1970. *Syllable, word, nexus, cursus*. The Hague: Mouton.
- Ranson, Diana. 2008. "La liaison variable dans un corpus du français méridional: L'importance relative de la fonction grammaticale". *Congrès mondial de linguistique française* ed. by Jacques Durand, Benoît Habert & Bernard Laks, 1657–1671. Paris: Institut de Linguistique Française. DOI: 10.1051/cmlf08279
- Schane, Scott A. 1965. *The phonological and morphological structure of French*. PhD Thesis. Boston: Massachusetts Institute of Technology.
- Tomasello, Michael. 2003. *Constructing a language: A usage based theory of language acquisition*. Cambridge: Harvard University Press.
- Tomasello, Michael. 2008. *Origins of human cognition*. Cambridge: MIT Press.
- Tranel, Bernard. 1995a. "Current issues in French phonology: Liaison and position theories". *The handbook of phonological theory* ed. by John A. Goldsmith, 798–816. Oxford: Blackwell. DOI: 10.1016/0024-3841(95)90104-3
- Tranel, Bernard. 1995b. "French final consonants and nonlinear phonology". *Lingua* 95. 131–167.

- Weinreich, Uriel, William Labov & Marvin I. Herzog. 1968. "Empirical foundations for a theory of language change". *Directions for historical linguistics* ed. by Winfred P. Lehmann and Yakov Malkiel, 95–188. Austin: University of Texas Press.
- Wimmer, Gejza & Gabriel Altmann. 1999. *Thesaurus of univariate discrete probability distributions*. Essen: Stamm Verlag.
- Zipf, George K. 1949. *Human behavior and the principle of least effort*. Boston: Addison-Wesley Press.

PART II

**Sources and functions
of sociophonetic variation**

Derhoticisation in Scottish English

A sociophonetic journey

Jane Stuart-Smith*, Eleanor Lawson** and James M. Scobbie**

* English Language / Glasgow University Laboratory of Phonetics,
University of Glasgow

** CASL Research Centre, Queen Margaret University Edinburgh

This paper presents the rewards of a sociophonetic journey by focusing on fine-grained variation in Scottish English coda /r/. We synthesize the results of some 15 years of research and provide a sociophonological account of variation and change in this feature. We summarize observations on coda /r/ in Scottish English across the twentieth century, which reveal a socially-constrained, long-term process of derhoticisation in working-class speech, alongside strengthening of /r/ in middle-class speakers. We then consider the linguistic and social factors involved, information from studies based on listener responses, the acoustics of derhoticisation, and insights gained from a socio-articulatory ultrasound corpus collected. These different views of coda /r/ force us to consider carefully the complex relationships between auditory, acoustic, and articulatory descriptions of (socially structured) speech. We conclude by discussing the implications of our results for mental representations of speech and social information for speaker-hearers in this community.

1. Introduction¹

This paper presents the concrete example of the rewards of a sociophonetic journey by focusing on an area which is particularly rich and informative – fine-grained variation in Scottish English coda /r/. We synthesize the results of some

1. Jane Stuart-Smith is grateful to the editors for their comments and patience, and to audiences at the Workshop on Sociophonetics (Pisa), and at seminars at the Universities of Oxford, Stanford, and Berkeley, for their feedback on earlier versions of this paper. The research presented here was supported by awards to Jane Stuart-Smith from the Leverhulme Trust (F/179/AX), the AHRC, and the ESRC (R000239757), and to James M. Scobbie and Jane Stuart-Smith, from the ESRC, RES-000-22-2032 and also to Eleanor Lawson, RES-062-23-3246. This article has associated sound files and ultrasound video files for Figures 6, 9, 11, 12 and 13. Please find them on <http://dx.doi.org/10.1075/silv.15.media>

15 years of research, including our current work in progress, with those of previous studies, and provide a sociophonological account of variation and change in this feature. This forces us to consider carefully the complex relationships between auditory, acoustic, and articulatory descriptions of (socially structured) speech. Our research also raises questions about speakers' mental representations of such information.

We begin by summarizing observations on coda /r/ in Scottish English across the twentieth century, which reveal a socially-constrained, long-term process of derhoticisation. Then we consider the most recent evidence for derhoticisation from different perspectives in order to learn more about the nature and mechanism of the change. We look at the linguistic and social factors involved (Sections 2 and 3); the views from the listener (Section 4); the acoustics of derhoticisation (Section 5); and insights from a socio-articulatory corpus collected and analysed using Ultrasound Tongue Imaging (Section 6). Finally we discuss the implications of our results for representation, by analysts, and for speaker-hearers in this community.

1.1 Derhoticisation in Scottish English in the twentieth century

Scottish English is a range of varieties forming a sociolinguistic continuum between two poles, broad vernacular Scots spoken by working-class speakers at one end, deriving historically from Northern forms of the Anglian dialect of Old English, and Standard Scottish English (SSE), spoken by middle-class speakers at the other, continuing varieties of Southern English which were adopted by the upper classes from the seventeenth century onwards, and later used increasingly by middle-class speakers (e.g. Stuart-Smith 2003; Durand 2004). In the conurbations of the Central Belt of Scotland stretching between Edinburgh and Glasgow (Figure 1), home to most of the population, many speakers drift up and down the continuum according to formality, context and interlocutor (Aitken 1984). In these urban areas, stratification by social class is still strongly adhered to at both ends of the continuum, with a continual process of social (and geographical) mobility in between (e.g. MacFarlane & Stuart-Smith 2012).

Accents of English which have a phonological specification of consonantal /r/ in coda position (also called 'postvocalic /r/') in words such as *car*, *card*, are often referred to as 'rhotic'. Scottish English is the classic rhotic variety of English in the UK (Wells 1982). Although /r/ was once an apical tap [ɾ] and often a trill [r] (Grant 1914; Johnston 1997), at least since the turn of the nineteenth century, derhoticisation in working-class speech, alongside an increasing use of approximant forms of /r/, have led to a sociophonetic continuum in the realization of postvocalic /r/.

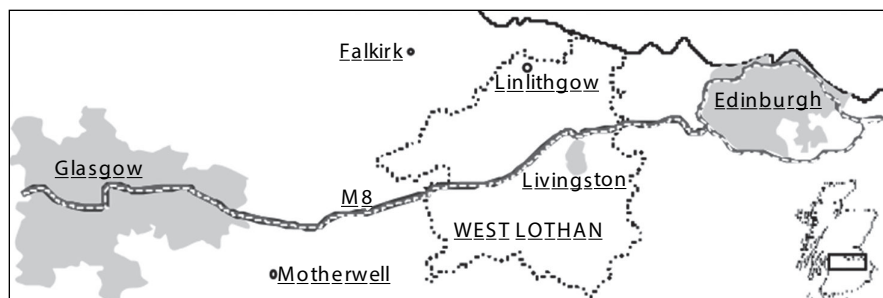


Figure 1. The Central Belt of Scotland (see inset) showing the cities of Glasgow on the west, Edinburgh on the East, and Livingston in between (from Lawson et al. 2008).

By derhoticisation, we mean either, diachronically, the gradient phonetic lenition process from trill towards a complete loss of /r/, or, synchronically, productions of /r/ weakly exhibiting few or none of the correlates typically attributed to its rhotic status. We survey the evidence for derhoticisation briefly below.

Reports of weak rhoticity in the realization of postvocalic /r/ date back to the early twentieth century, when reports of accent variation are first available. They relate to Scottish English spoken on the West coast, and specifically, as characteristic of the urban speech of the ‘degenerate Glasgow-Irish’, to whom numerous undesirable speech and language habits were attributed, including the infamous glottal stop (Trotter 1901 in Johnston 1997: 511). Polite speakers were noted to use the apical trill [r] or tap [ɾ] (Williams 1909; Grant 1914), or the postalveolar approximant [ɹ] (though, at this point, approximant /r/ was not considered a ‘Scottish sound’ by Grant and Dixon 1921, in Romaine 1978). All these realizations are attested in the very short reading passages recorded by William Dögen for the Berliner Lautarchiv in 1916/17 from young male speakers from Glasgow and surrounding areas (Richmond 2013). By 1938, approximant [ɹ] was a recommended realization for the ‘student of good speech’, as acceptable as [ɾ], and more so if speakers wished to achieve the socially more desirable merger of /ʌ ɪ ε/ to /ɜ/ in a prerhotic context, e.g. in the words *fur*, *first* and *herb* (McAllister 1938; Lawson et al. 2013, forthcoming).

The earliest indication of derhoticisation in Edinburgh is indirect, from observations made in the Edinburgh Articulation Test (EAT), a standardized study of articulation in children’s speech aged 3.0 to 5.6 carried out in the late 1960s. The authors of the EAT coded vocalized variants along with consonantal /r/, stating: “many Scottish 2½-year-old children used a diphthong in positions where they later developed one of the many forms of [r]. *As this diphthong may also be an acceptable adult realisation, it had to be considered correct in this context.*” (Anthony

et al. 1971:6, in Scobbie et al. 2007, their emphasis). Note that such diphthongs may also have been picking up recessive upper-middle-class non-rhoticity.

A clearer picture of derhoticisation in Edinburgh is made possible thanks to two early sociolinguistic studies, carried out by Romaine (1978) and Johnston (Speitel & Johnston 1983; Johnston 1985). Romaine's study concentrated on working-class children. Her results showed that boys were less rhotic than the girls, who also used more instances of postalveolar [ɹ], as opposed to tapped or trilled variants. Non-rhoticity was also more common in the wordlists than in spontaneous speech. Romaine interpreted the non-rhoticity in the boys as a vernacular change from below taking place in Scots, "which happens to coincide with a much larger national norm" (i.e. 'national' in a UK sense, indicating non-rhoticity in RP, p. 155). She saw non-rhoticity as carrying covert prestige, and part of a local system of differentiation from the more socially-desirable postalveolar approximant [ɹ] favoured by the girls, associated with middle-class speakers and prestigious varieties of Highland English (p. 156).

Johnston's study worked with a much larger socially-stratified corpus of adults. He observed two very different kinds of non-rhoticity: that found in older (55–79 year old) upper middle-class women, and that at the opposite end of the social-gender continuum, lower working-class men (18–55 years old), who showed vocalization to a 'strongly pharyngealized vowel'. Such an outcome is not surprising since articulated /r/ in this speaker group is typically 'dark', with secondary pharyngealization. Johnston also found that in coda position, postalveolar [ɹ] was favoured particularly by younger female speakers, and in more formal styles. He suggested that postalveolar [ɹ] was "a recent innovation, probably from middle-class RP, into Edinburgh speech" (p. 27). Johnston interpreted the motivations for both changes in terms of the social dynamics within Scotland. Derhoticisation was identified as showing 'street-smart' associations; rhoticity in the middle classes was seen as reflecting constructions of a resurgence of Scottish identity in the Scottish middle-classes, expressed in a 'home-grown model of Standard Scottish English' used in preference to, and a reaction against, earlier local Scottish prestige models close to RP.

Back on the West Coast, Macafee's (1983:32) description of Glaswegian dialect, outlined similar derhoticisation to plain or pharygealised vowels in working-class speakers. Subsequent quantitative analysis of a socially-stratified corpus of Glaswegian collected in 1997 confirmed substantial derhoticisation in working-class speakers, especially adolescents (Stuart-Smith 2003; Stuart-Smith et al. 2007). Derhoticised reflexes fell into two main categories: pharygealised/uvularised vowels, favoured by boys in a specific phonological context (before a consonant, e.g. *card*); and plain vowels with no audible secondary 'colouring', favoured by girls in unstressed prepausal position, e.g. *better*#, though both groups showed numerous

instances of both variants. Middle-class speakers tended to be rhotic, with both older and younger speakers favouring postalveolar and/or retroflex approximants, especially younger middle-class girls. (If articulatory /r/ was produced by working-class speakers, it was usually a tap.)

Overall, the evidence for the twentieth century suggests the development of a socially-stratified rhotic-derhotic continuum in the Scottish English of the Central Belt, with weakly articulated, or vocalized, rhotics in working-class speech contrasting with audibly strong rhotic approximants in the aspiring middle-classes. We now turn to the sociolinguistic evidence for the progress of derhoticisation, and the corresponding development of the continuum, in the early 21st century.

2. Derhoticisation in Scottish English in the 2000s

In 2003, a further corpus of Glaswegian was collected from an age-stratified sample of working-class speakers from the same area as the 1997 corpus (e.g. Stuart-Smith 2006; Stuart-Smith & Timmins 2010). Figure 2 shows the substantial derhoticisation that was found in these speakers. Like Romaine (1978), derhoticisation was more prevalent in read wordlists. This stylistic shift away from the regional standard norm (rhoticity) in a reading task confirms that this feature still carries the kind of covert prestige suggested by Johnston.

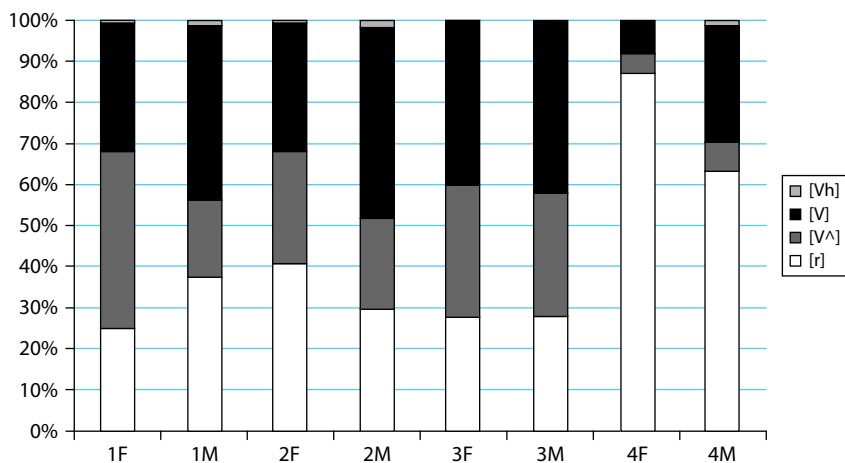


Figure 2. Distribution of variants of postvocalic /r/ in 48 speakers of Glaswegian in 2003, n = 1889. M = male, F = female; 1 = 10–11 years; 2 = 12–13 years; 3 = 14–15 years; 4 = 40–60 years. [r] = articulated variants of /r/; [V^] = vowels with audible pharyngealisation/uvularisation; [V] = plain vowel; [Vh] = vowel followed by audible frication.

Only six years had elapsed by the time we collected the 2003 corpus, so it is difficult to know the extent to which variation over this time reflects real-time change (Labov 1994). Comparison of the percentage of use of the plain vowel variant for coda /r/ for individual speakers in 1997 (8 children) with those recorded in 2003 (36 children) suggests that derhoticisation is a very gradual change in progress. The speakers from 1997, shown as dark bars, fit within the distribution of the speakers from 2003; see Figure 3.

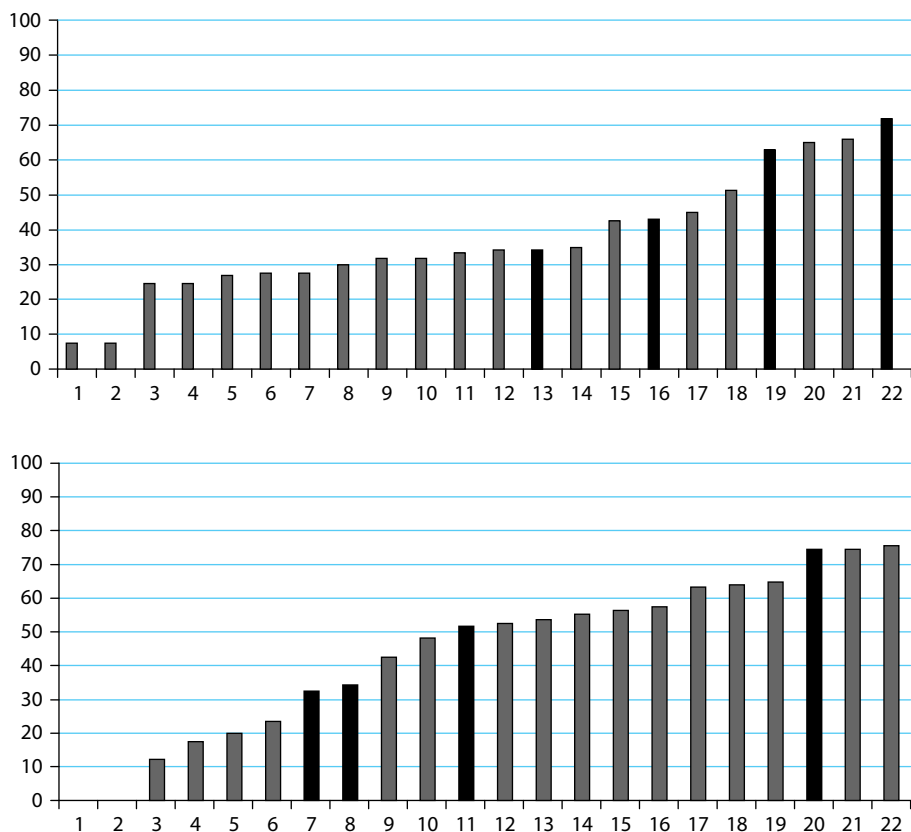


Figure 3. Percentage of the plain vowel variant for coda /r/ used by 42 speakers, 36 recorded in 2003 (pale bars) and 8 recorded in 1997 (dark bars). The first chart shows female speakers, the second male speakers.

Previous studies had concentrated on the two cities at either end of the Central Belt. In 2007 a corpus of speech and articulatory data (tongue movement) was collected from working-class adolescents in Livingston, a new town, in between, but lying closer to Edinburgh than Glasgow (Figure 1); Lawson et al. (2008). Auditory transcription showed some derhoticisation, but on average only 20% of

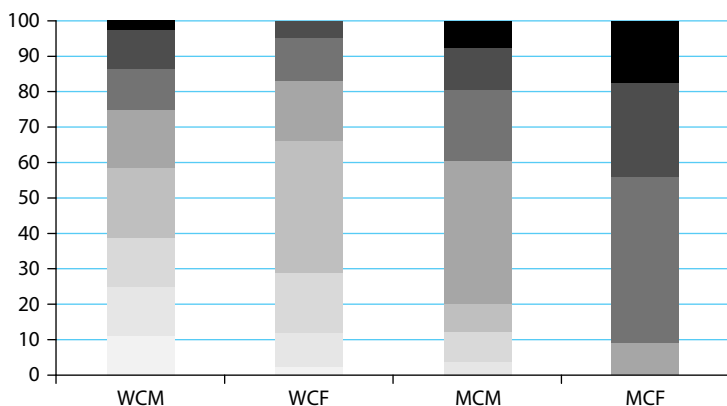


Figure 4. Bar graph showing the percentage of auditory variants used by each socioeconomic and gender group in the ECB08 corpus. WC/MC = working/middle-class; M/F = male/female. Paler grey segments represent rless and weakly rhotic variants, while darker grey segments represent strongly rhotic variants. N = 139. From Lawson et al. (2011), Figure 2.

all postvocalic /r/, which is considerably less than the amount found in Glasgow. Also unlike Glasgow, the most common environment for non-rhotic tokens was in stressed syllables in utterance final position (e.g. *car##*), though the next most likely context was in unstressed syllables in utterance final position (e.g. *better#*).

A year later, a socially-stratified audio and articulatory corpus (ECB08) was collected from middle-class adolescents in Edinburgh, and working-class adolescents again from Livingston. The study was designed to further explore possible articulatory mechanisms for derhoticisation. The auditory assessment of postvocalic /r/ drawn from the wordlist confirms more weakly-articulated /r/ and derhoticisation (pale grey segments) in working-class speakers towards the East, and illustrates well how the rhotic-derhotic continuum is constrained by social class and gender (Figure 4; Lawson et al. 2011).

It is clear that these recent data continue the earlier trends. Middle-class, and especially female, speakers are leading a change from above towards audibly 'strong' approximant /r/. These changes exploiting the variant [ɹ], which may be of Anglo-English origin, to mark both more confidence in a specifically *Scottish* (not UK) middle-class identity (Johnston 1985), and social differentiation from Scottish working-class identities (Douglas 2009). Working-class speakers on the other hand are participating in long-term vernacular change from below, resulting perhaps in the completion of derhoticisation which will be non-rhoticity. The earliest reports pin the latter change to the turn of the twentieth century, but the change may have started much earlier. The progress of derhoticisation varies according to location, but is more advanced in the more populous western conurbation.

Another important aspect of Scottish derhoticisation is how it relates to non-rhoticity in English English. For it cannot be ignored that in some phonetic contexts, e.g. following /a/, the derhoticised reflexes in Glasgow appear strikingly non-rhotic, making the outcome phonetically very similar to the non-rhoticity found in the UK standard (and indeed non-standard) varieties of English English (Romaine 1978). Moreover, the recent large-scale study of rhoticity along the Scottish-English Border has also found derhoticisation in younger speakers, though with significantly more at the western end (Gretna) than in the more Scottish, east-coast, town of Eyemouth, which aligns with attitudes of Scottishness (Llamas 2010). Pukli and Jauriberry (2011) also report some derhoticisation in the rural south-western city of Ayr, as well as the substantial appearance of postalveolar [ɹ] in onset position, and more generally in young female speakers. Just as other consonantal changes appear to be making their way north (e.g. TH-fronting, L-vocalisation; Stuart-Smith et al. 2007), there is a possibility that the Glaswegian non-rhotic outcome could also reflect the effective confluence of two streams of change, one a vernacular change within Scots, and the other a contact-induced change from non-rhotic varieties of English. In order to consider the empirical evidence for this, in the following section we put derhoticisation in the context of the wider system of changes in progress in Glaswegian, and the social factors which are involved in their transmission.

3. Social factors in Glaswegian derhoticisation

The most recent study of derhoticisation of /r/ in Glasgow was undertaken as part of a broader variationist project. Its aim was to consider the role of a large range of social factors in several sound changes in progress in Glaswegian, including opportunity for contact with speakers of dialects furth of the city, and the possible influence of the broadcast media. Also in the 1997 corpus, derhoticisation of postvocalic /r/ was found in the speech of those working-class adolescents who were leading in the rapid adoption of some consonant features typically associated with London and southern English, specifically the use of [f] and [v] for /θ/ and /ð/ (TH-/DH-fronting), and vocalization of coda /l/ to a high back (un)rounded vowel (L-vocalisation). That these speakers were also the least geographically and socially mobile posed a challenge for contact-based theories of the diffusion of these changes (e.g. Trudgill 1986), and the media themselves suggested that watching television, and in particular, dramas set in London, like the exceptionally popular soap, *EastEnders*, was a key factor.

The Glasgow Media Project constituted the first comprehensive systematic sociolinguistic investigation of the influence of the broadcast media on language change, by focusing on the possible role of exposure to, and psychological engagement with, London-based TV dramas on Glaswegian vernacular phonology. Three groups of linguistic variables were considered:

- consonant innovations: e.g. TH-fronting. Three rapid changes in Glaswegian look like instances of diffusion from Southern varieties of English English, which took off in the 1990s, though they are sporadically reported in Scottish English much earlier (Macafee 1983; Anthony et al. 1971);
- ongoing vernacular changes: e.g. derhoticisation of postvocalic /r/. As noted before, this change appears to be system-internal, though the final outcome (e.g. non-rhoticity) can coincide phonetically with English English norms;
- more stable sociolinguistic variation: e.g. realization of the vowels /a/, /u/ and /ɪ/.²

Only the consonant innovations have been explicitly linked with exposure to London English on the television. However to test the hypothesis that television might be a contributory factor in the innovative changes, we needed also to test those variables for which media influence has never been mooted, and so vowels and derhoticisation were included in the study.

The auditory variants for the consonant innovations (e.g. [f], [v]), and derhoticisation (/Vr/ sequences realized as a plain vowel with no velar or pharyngeal quality), and F1 and F2 of /a u ɪ/, for read (wordlists) and spontaneous (conversational) speech were the dependent variables in a series of regression models constructed for the 36 adolescent informants. The independent variables consisted of representative linguistic factors (e.g. position in the word, adjacent phonetic context), and a large array of extralinguistic factors: opportunities for dialect contact with speakers of other English dialects; attitudes to dialects elicited from responses to audio recordings and paper surveys; engagement and participation in a range of social practices; preferences for music and radio, film (cinema, DVD, video); activity on the internet and engagement with computer games and computer-mediated communication; and exposure to, and psychological engagement with, the television. The variables were drawn from a structured questionnaire completed by each informant, an informal interview with the fieldworker, their own spontaneous speech recordings with their friends, and participant observation by the fieldworker during the period of data collection. Full details and results of the regression study can be found in Stuart-Smith et al. (2013).

2. That was our hypothesis at the time. In fact the new Glasgow Real-Time Project is demonstrating real-time change in /u/ (e.g. Rathcke et al. 2012).

The main findings were:

- The consonant innovations were strongly constrained by linguistic factors and by several extralinguistic factors including: participation in anti-school social practices, such as adopting Glasgow street style in place of school uniform; strong psychological and emotional engagement with the London-based TV soap opera, *EastEnders*; reported contact with friends and relatives in England; and more weakly, with positive attitudes to London accents.
- The vowel variables showed almost exclusively significant effects for linguistic factors, with very little evidence for social factors of any kind.
- The results for derhoticisation of postvocalic /r/ were split according to speech style. (i) In spontaneous speech derhoticisation patterned like the vowel variables: the predominant effects were for the linguistic factors, with very little evidence for social factors. (ii) In read speech derhoticisation showed a similar pattern to the consonant innovations: both linguistic and social factors were significantly correlated. Plain vowels for /Vr/ were more likely in unstressed prepausal position (e.g. *better*#), and they were linked with anti-school social practices, strong psychological engagement with *EastEnders*, and the ability to correctly identify their own local accent from a recording, amongst other factors (including participation in sport and playing football); only dialect contact proved to be consistently non significant.

To summarize: there is no evidence that direct contact with non-rhotic English speakers promotes derhoticisation. But indirect contact with non-rhotic London English, by psychologically engaging with a TV drama set in London, does seem to be a factor, but only for a particular speech style, reading a list of words out loud.

The results for derhoticisation indicate that the change is not entirely driven by system internal forces. At the same time, they contribute to our understanding of media influence on speech more generally. The evidence from Glasgow shows that only some phonological features are linked to engagement with the television. This supports an extension of existing models of media influence in mass communications theory to language, specifically that speaker/viewers use their social and linguistic knowledge to ‘decode’ televised speech, so here Glaswegians parse *EastEnders* through the filter of their own experiences as active members of actual speech communities within the city (Hall 1980; Gunter 2000; Stuart-Smith 2011). The assumption is that viewers largely filter out aspects of media language which are irrelevant in terms of social meaning and linguistic structure (which is probably the majority of most experienced media material). But sometimes a viewer’s existing features may be enhanced provided that there are points of reciprocity and alignment with the viewer’s own local social context and linguistic system (which is probably quite rare). So the consonant innovations look like diffusing

features ‘from outside’ the dialect, hopping north from London. While there is some support for dialect contact being involved, closer up they look fundamentally like local system-internal variation which is, as it were, bubbling up, developing a variety of social symbolic functions, which in turn speed up the changes in progress (Eckert 2000). Media influence represents an additional factor through which speakers enhance their existing variation, thus fuelling their rapid acceleration through the system and the community.

Derhoticisation has been underway for many decades in Scottish English, apparently without influence of English English. Only in read speech is derhoticisation linked with indirect contact with London English via the TV. This helps unpick the processes of media influence further. When we recorded the working-class adolescents reading the wordlists, rather than read them ‘correctly’ (i.e. approximate Standard Scottish English, e.g. Labov 1972), our informants produced distinctly non-standard variants. Overall a specific position, or stance towards the task and fieldworker was taken (Jaffe 2009), as if distancing themselves and their speech from the University. The wordlists were rattled off, punctuated with laughter; they were highly performative, in the sense of Baumann’s construction for an audience (Coupland 2007). In terms of variation, the wordlists showed increased use of consonant innovations, and more derhoticisation. Previous research on stance-taking through language has noted that media representations can simplify social-indexical relationships, and so speed up linguistic appropriation from the media (see e.g. the spread of the catchphrase ‘Whassup?’ in American English; Bucholtz 2009:288). Aspects of language which index nuances of interpersonal interaction, and subsequently local micro-social relationships, can then be used in the media, e.g. in advertising, with much broader indexical referents.

We hypothesize that stance, and/or other kinds of social informativity of linguistic variation (Pierrehumbert 2006; Eckert 2008), may be a determining factor in whether speaker/viewers’ sociolinguistic systems may respond to media language. The enhancement of viewers’ existing features may depend on the implicit recognition, or mapping, of linguistic features indexing particular stances in media language, with the possible indexing of stancetaking in their own interactions. Crucially, being perceivers *and* producers of social variation, or being listeners using their ‘speaking brain’ (Keith Johnson personal communication), is also important here; Kuhl, e.g. (2010). The interesting point about the link between engagement with the TV and derhoticisation is of course that this change has never been interpreted as a contact-induced change. These results emphasize the importance of the speaker/viewer’s local social-phonological system in the decoding of televised speech. They also suggest mechanisms for how existing local variation could become accelerated through indirect contact with accent features, albeit through strong psychological and emotional engagement with a television

programme and its characters. We suspect that direct contact with English English does not emerge as a factor precisely because this is mediated by ideological and attitudinal factors relating to nationality and non-rhoticity (Llamas et al. 2009).

Thus, teasing apart the social factors that contribute to the progress of derhoticisation is both informative in understanding the change itself, and for modeling media influence on speech. There is indeed some evidence to support the view that non-rhoticity in western Central Scottish English reflects the outcome of two streams of change, though the nature of the contact-induced change needs to be refined to indirect contact with non-standard English via the broadcast media. But we still need to discover the phonetic mechanisms underpinning derhoticisation, and the rhotic-derhotic continuum; in order to do this we must consider the phonetic data – and how they might be represented.

4. Scottish derhoticisation and the listener

The variation observed in the Scottish English rhotic-derhotic continuum, provokes two challenging questions: (1) What is the phonetic nature of the derhoticised reflexes? (2) How can we best capture this complexity. Until recently representing sociophonetic variation was limited to characterizing aspects of the recorded speech signal, by auditory or acoustic analysis. Whilst it is increasingly assumed that acoustic analyses are superior to auditory ones, and certainly they have the advantage of yielding continuous measures which are amenable to more robust statistical analyses (e.g. Warren & Hay 2012), both are equally valid. Each gives a different (and incomplete) picture of the ‘same’ thing; both are connected, but not in straightforward ways, and in turn make inferences about underlying articulatory gestures. In this, and the next two sections we review previous and ongoing phonetic work on derhoticisation which exemplifies these points. We begin by considering the view from the listener, both the analyst and the speech community. In Section 5 we shift perspective to look at acoustic representations. In Section 6 we move closer to articulation, using Ultrasound Tongue Imaging.

4.1 The listener as analyst: Auditory phonetic representations of derhoticisation

All the studies discussed above used impressionistic or auditory transcription. Using this method, analysts categorise the auditory continuum of variation in ‘articulatory’ terms, i.e. the analyst constructs a kinaesthetic interpretation of the possible articulatory strategy used by the speaker, and then represents it using

IPA symbols (Ogden 2009). Transcription can be more or less detailed, but usually results in fairly broad, discrete categories, which make strong assumptions about the articulatory gestures underlying the auditory objects. Whilst auditory transcription is a valid and useful method of representing phonetic variation, we need to be mindful that it yields auditory, not articulatory, objects. It also requires the analyst to broadly divide up and assign parts of the auditory continuum to one or other categories, whereas listeners may feel that aspects of more than one category may be involved. Social-indexical ingrained variation may not be easily audible even to trained phoneticians (Docherty & Foulkes 1999).

Each group transcribing derhoticisation came up with different solutions,³ which in turn coloured their theoretical perspective. For example, recognizing many possible variants emphasizes gradient progression of the change, as opposed to coding with or without final /r/, which points to the final outcome (contrast 'derhoticisation'/'R-vocalization' with 'R-Loss'). For all, the transcription of the derhoticised variation was extremely difficult, and this motivated a small-scale study to investigate this analytical task (Stuart-Smith 2007).

A subset of the 2003 Glasgow corpus was selected, 12 male working-class informants, nine adolescents, with three from each age group, and three adults. All the adolescents were observed to show derhoticisation in the main study. A subset of words were selected from the larger wordlist, in which /r/ follows the low vowel /a/: *heart, barn, farm, car, far, card*. These were subjected to a narrow auditory phonetic transcription by three phonetically trained transcribers: 1: CT, a Scottish-English, rhotic middle-class speaker from Edinburgh; 2: JSS, an English-English, non-rhotic middle-class speaker from Southern England; 3: RL, a Scottish-English, rhotic middle-class speaker from a small town just south of Glasgow. The results of the transcriptions are shown in Figure 5.

The results are striking. Each transcriber hears the same signal, but transcribes and categorises it differently from each other (see also Plug & Ogden 2003). All heard some derhoticisation, CT the least, and RL the most – so interestingly the outcome is not straightforwardly predicated on the transcriber being rhotic (Yaeger-Dror et al. 2009), though perhaps differential experience of the rhotic-derhotic continuum, and/or the socially symbolic nature of derhoticised variants might play a role. Recall that derhoticisation is more advanced on the West than

3. Speitel & Johnston (1983) and Stuart-Smith (e.g. 2003) used very narrow auditory phonetic transcription and identified a range of different kinds of derhoticised and/or vocalic outcome, which can be represented either as extremely weak uvular approximants, or vowels with secondary pharygealisation. Romaine recognized this phonetic complexity but opted to represent a simplified set of categories, grouping plain and coloured vowels together as complete deletion. Lawson et al. (2008) simply divided variants into rhotic and non-rhotic.

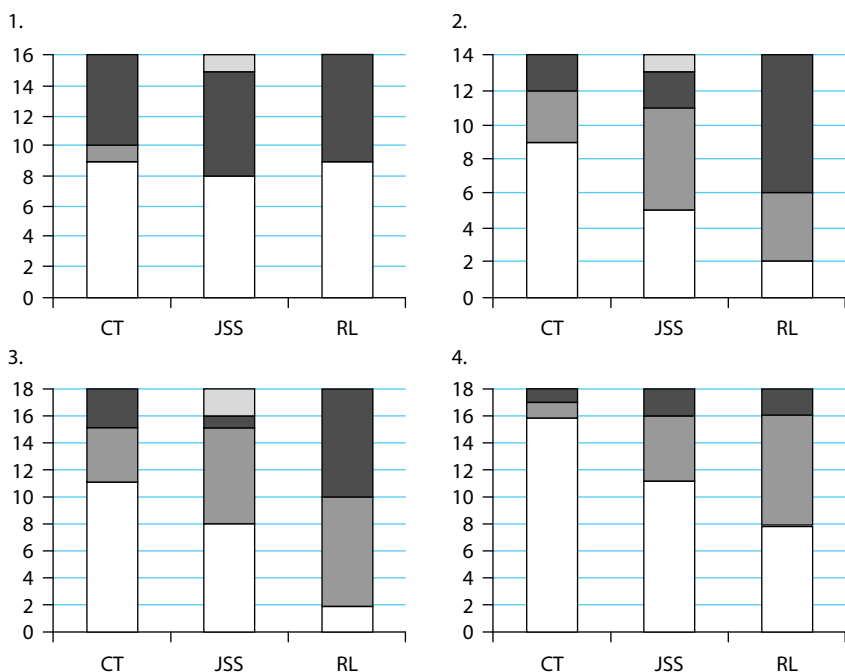


Figure 5. Results of the auditory transcription of postvocalic /r/ in word-list data read by 12 male Glaswegian working-class speakers, organised into four age groups (1 = 10–11, 2 = 12–13, 3 = 14–15, 4 = 40–60). The judgments of the three transcribers (CT, JSS, RL) are shown in each chart from left to right. White = articulated /r/, spotted = pharyngealised/uvularised vowels, grey = plain vowels, striped = vowels followed by [h] or [ɦ], from Stuart-Smith (2007: 1308).

the East. We also found that whilst transcribers effectively segmented the auditory continuum at different points, they were internally consistent.

There is also another key shared feature. All three transcribers found that they could not assign what they heard only to two categories, ‘plain vowel’ or some kind of articulated /r/ (the phonetic variants are grouped together for this representation but ranged from weak approximants to weak taps). A third auditory category was needed for variants which fell between articulated /r/ and no audible articulation at all, which could be termed either as ‘extremely weak uvular approximants’ or as ‘pharyngealized or uvularized vowels’. This could be interpreted in a prescriptive way as analysts simply being unable to implement the IPA categories appropriately. But we will see that the acoustic, and especially the articulatory, data show that a category to accommodate such a variable percept – hearing sometimes a consonantal gesture and sometimes not – is well motivated.

4.2 The listener in the community: Evidence from speech perception

An alternative view to that of the analyst can be drawn from perceptual evidence from the community – how listeners parse, and/or respond to variants along the rhotic-derhotic continuum. Carey (2010) carried out a small-scale study of cross-linguistic dialect perception, looking at Glaswegian and Southern British English (SBE) listeners' responses to stimuli from both dialects. Judges listened to three pairs of sentences which varied according to whether postvocalic /r/ was present or absent, e.g. *That surprise for the child* vs *That's a prize for the child*, or *The congregation certainly likes arms* vs *The congregation certainly likes psalms*, and then had to write down what they heard (the stimuli examined a large number of phonological differences between SBE and Glaswegian). Glaswegian listeners found it as difficult as SBE listeners to recover postvocalic /r/ in such sequences, even in the stressed monosyllable *arms*.

MacFarlane & Stuart-Smith (2012)'s matched guise study considered social evaluation. The same talker produced recordings of pairs of words which varied in the realization of a single variable. Listeners were led to believe that two speakers, Lee ('regular Glasgow') and Phil ('socially-aspirational Glasgow Uni(versity)') had produced the recordings, and were given only a group of brand logos for each 'speaker' as their guide to the lifestyles of the 'two' men. Three out of the four experimental variables related to /r/. The realization of onset /r/; the duration and quality of the final syllable of disyllabic words such as *number* (longer for Glasgow Uni, shorter and less rhotic for regular Glasgow); and the quality of the prerhotic vowel in words like *nerve* and *pearl* ([ə] is associated with Glasgow Uni – and also with vocalic rhoticisation; [ɛ] is associated with a following tapped /r/ variant and Regular Glasgow). Listeners were very good at correctly socially categorizing the 'talkers' using the *number* and *nerve* variables, i.e. the two variables which related to realization of postvocalic /r/. But the realization of onset /r/ was only categorized at chance level, refuting the hypothesis that taps in this position associate more with 'regular Glaswegian' speech, that is, working-class Glaswegian speech.⁴

4. This last result is intriguing since it suggests that the realization of coda /r/ carries more meaning for these speakers, than that of onset /r/. If this is right, this might also account for Johnston's suggestion that postalveolar approximant /r/ spread from English English into Edinburgh English in onset position. Pukli & Jauriberry's (2011:88) findings from Ayr that onset /r/ is increasingly being realized by an alveolar approximant [ɹ] in Ayr are also congruent. So too are the similar shifts observed at the western end of the Scottish/English border by Llamas 2010. The originally 'English' variant may have slipped more easily into the array of /r/ variation in this environment, becoming Scottish, but unmarked as such, precisely because variation in onset /r/ does less social 'work' than coda /r/. Our current work on articulation of /r/ is interrogating this assumption further.

These two studies both show that derhoticisation is also taking place perceptually for members of the community, and is not only restricted to the domain of the analyst. Both ends of the rhotic-derhotic continuum also still seem to carry the kind of locally-salient social meanings that were proposed by Johnston for Edinburgh. But if we want to pin down what listeners are responding to, it is clear that we need to go further than the admittedly tricky auditory categorization. Our next attempt was acoustic analysis.

5. The acoustic characteristics of derhoticising /r/

The difficulties with auditory percepts which were challenging to auditorily categorise, and themselves variable, motivated an acoustic analysis of the data whose auditory transcription was discussed above in §4.1 (Stuart-Smith 2007). Since it was also unclear whether the final outcome of derhoticisation to plain vowels is leading to a merger (recall that weakened /r/ is now perceptually variable even to Scottish listeners, §4.2), we included minimal pairs. To recap, we considered the acoustics of coda /r/ in the following words: *heart/hat*, *barn/ban*, *farm/fan* and *car*, *far*, *card*, in 12 working-class speakers, nine boys and three men.

We carried out a qualitative visual analysis of the spectrograms, and then used a parametric analysis of acoustic properties of the syllable rime, so e.g. *c-ar*, following the successful application of this method to variation in postvocalic /r/ in Dutch (Plug & Odgen 2003). This also addresses the practical difficulties of segmenting final /r/ consonants which were effectively no longer there. Using Praat, we labelled the waveform for the beginning and end of the vocalic portion (i.e. the entire duration of the vowel + /r/ portion of the syllable rime), and then measured the duration of the vocalic portion, the vowel quality in terms of the first three formants at the midpoint, and vowel tracks for the last five glottal pulses, again for the first three formants. Formant measures were extracted using Praat, and then corrected by hand.

The classic acoustic ‘signature’ of approximant /r/s, and also some trills and taps, is a lowered third formant (Lawson et al. 2011a; though see Heselwood & Plug 2011). The lowered F3 relates to the dimensions of a large cavity in the front of the vocal tract arising from specific articulatory gestures. The rather different configuration for uvular /r/ shows a different pattern of high and/or raised F3. Visual inspection of the spectrograms provided the following acoustic information for the four auditory variant categories shown above in Figure 5:

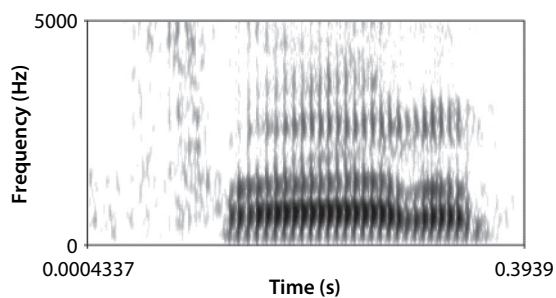
- articulated /r/: This included taps, a few weak approximants, and a single trill in the oldest man. The taps showed the expected momentary reduction in amplitude across the frequency range (Figure 6a), and the trill had four such

dips visible, reflecting four short interruptions in airflow (Figure 6b). In the few tokens of /r/ which were heard as (weakly) articulated approximant /r/, it is just possible to see the faint trace of the third formant dropping towards the end of the word, though just as striking is the reduction of amplitude above F2 (see Figure 6c).

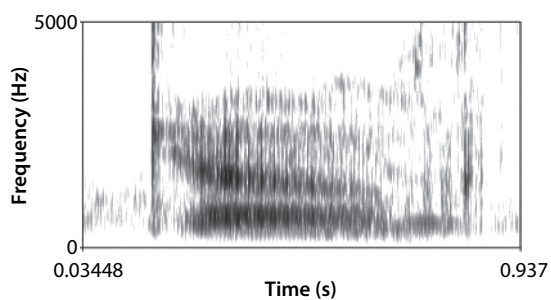
The other three variant categories capture different stages of audible erosion of the rhotic consonant:

- pharyngealized/uvularized vowel: These variants sound like extremely weak uvular approximants, or vowels with pharyngealization/uvularization. The primary acoustic characteristic is reduction in amplitude where /r/ would be expected. The weakened F3 is either flat or rising slightly (see Figure 6d);
- plain vowel: No primary or secondary articulation for a rhotic consonant was audible. The spectrograms typically show flat first and second formants, with very little energy above F2 (see Figure 6e). Inspection of successive spectra shows a very weak third formant which rises towards the end of the vocalic portion and into the voiceless period;
- vowel followed by audible frication: A small number of plain vowels sounded as if they were followed by a very weak fricative, possibly glottal, pharyngeal or even uvular. In Figure 6f, the vowel gives way to a period of very weak energy, with initial energy loss in F3, and then voicing ceases, though a period of very weak aperiodic noise is still visible for several ms.

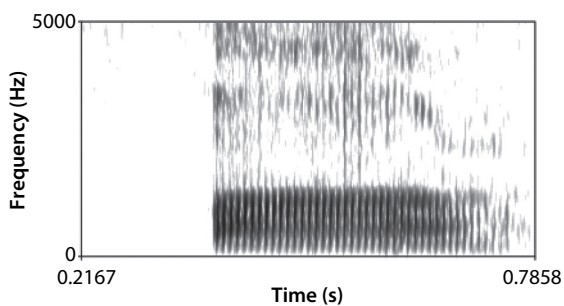
Neither first and second formant measures, nor durations, differed according to whether an articulated /r/ was audibly present or absent. Derhoticisation is not reliably distinguished through these measures. On the whole F3 was very difficult to measure because – as was observed – towards the end of the vocalic portion, where an acoustic reflection of the /r/ sound might be expected to be seen, there was a sharp drop in intensity in and above the region of F2, and in the F3 region. If it was possible to pick out F3 in speakers whose variants were audibly less rhotic, F3 was either flat or rising slightly, consistent with uvularization. This is illustrated in a comparison of the formant tracks from the most audibly rhotic boy with his much less rhotic-sounding friend (see Figure 7). A further result is that the derhoticized outcomes of /r/, even plain vowels, are still significantly distinct from words without <r>, so e.g. derhotic *heart* shows a longer, more retracted vowel than *hat*. This suggests that, at least for wordlist data, there is not yet a loss of phonological /r/, which is hinted at by Carey's (2010) results; it is likely that as in other non-rhotic varieties of English, the contrast will be maintained by differences in the vowel system (for further discussion of the impact of rhoticity on Scottish vowels, see Lawson et al. 2013).



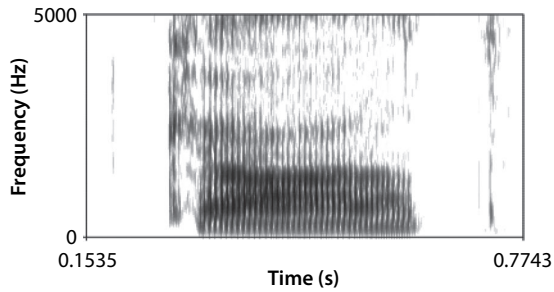
a. *farm* with tap (adult male)



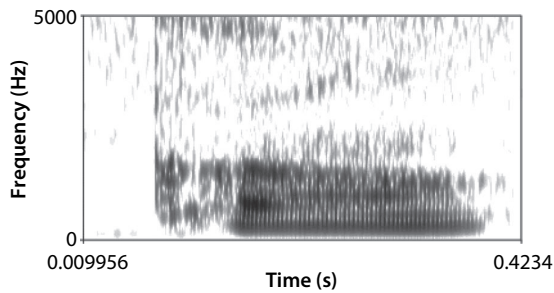
b. *car* with trill (adult male)



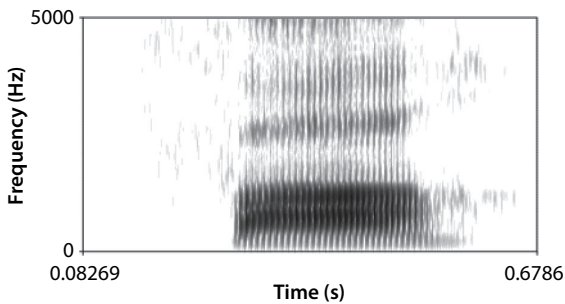
c. *far* with weak approximant (14 yr-old boy)



d. *card* with pharyngealized/uvularized vowel (12 yr-old boy)



e. *car* with plain vowel (11 yr-old boy)



f. *far* with vowel followed by weak frication (14 yr-old boy)

Figure 6. Spectrograms illustrating the four auditory variant categories shown in Figure 5. Articulated /r/ is shown on the left in (a)–(c); vowel variants on the right – pharyngealized/uvularized vowel (d), plain vowel (e), and vowel followed by weak frication (f). All recordings were made in 2003 in Glasgow.

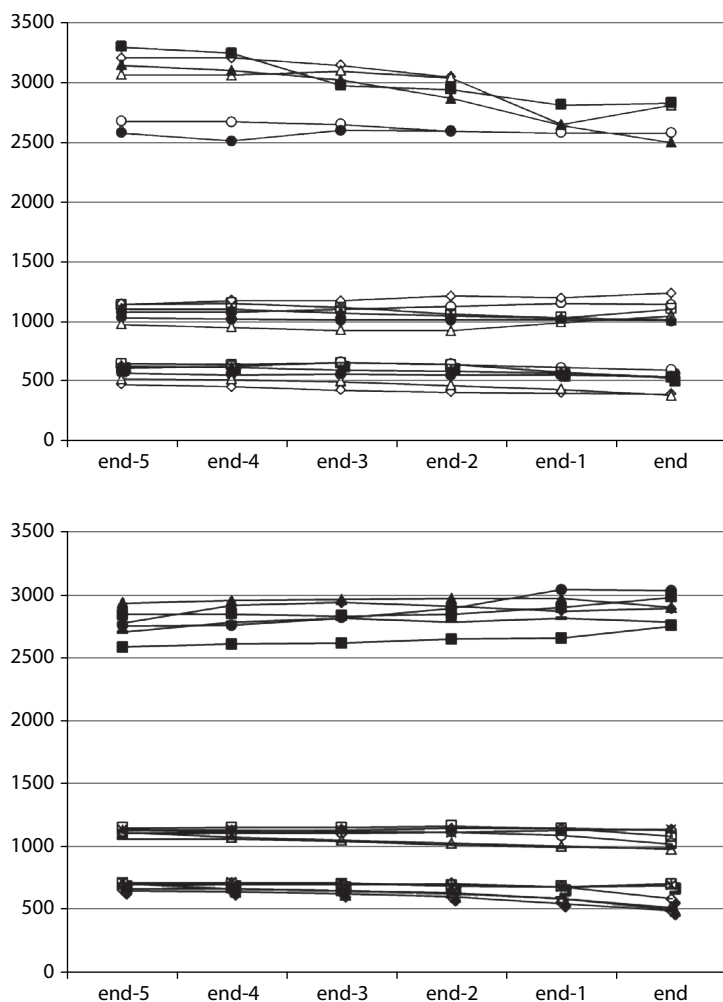


Figure 7. Handcorrected time-normalized formant tracks taken at the end of the vocalic portion and for each of the five preceding pulses, for the first three formants for two speakers: (a) 14 year-old boy heard as rhotic, shows slight dip in a high F3 in most words with /r/ (this boy produced *far*, Figure 6c). (b) 14 year-old boy heard with mainly pharyngealized vowels for words with /r/, shows high, flat or rising F3, with weak amplitude (this boy produced *far*, Figure 6f).

The outcome of the acoustic analysis is not as helpful as we had hoped. In part this is because the reflexes of derhoticisation do not relate easily to known acoustic parameters. Rather the clearest common characteristic is a reduction of acoustic energy above F2. On the one hand, these stretches of very weak formant energy, with and without, voicing, may help account for the variable auditory percepts of

rhoticity. That is to say, sometimes there is, and sometimes there is not, some kind of secondary pharyngeal articulation, and so the residue of an articulated /r/ is still present. But on the other, lack of energy in a specific frequency region makes it difficult to identify and measure formants in and above that region. Quantitatively capturing such acoustic weakening itself is also far from straightforward. This reminds us that acoustic analysis may not always be superior to auditory analysis; it is necessarily partial and more subjective than it might appear (Ogden 2009: 36). Thus the acoustic analysis moves us forward, but it still leaves us with another picture of the data, as opposed to a better understanding of the mechanism of derhoticisation.⁵ For this we need to turn to articulatory views of the phenomenon.

6. Investigating derhoticisation using articulatory data

Auditory-acoustic challenges led us to consider a different kind of phonetic representation, closer to the articulatory strategies involved, achieved using Ultrasound Tongue Imaging (UTI), and arising from a 2004 study of Dutch /r/ (Scobbie & Sebrechts 2011). Our Scottish work is in progress, and in this section we report some key relevant findings from three recent studies carried out on the Eastern Central Belt, a small pilot reported in Scobbie (2007), a sub-project to assess the feasibility of UTI for sociolinguistic fieldwork (WL07 corpus from Livingston; Lawson et al. 2008), and a socially-stratified articulatory speech corpus, with middle-class speakers from Edinburgh and working-class speakers from Livingston (ECB08 corpus; e.g. Lawson et al. 2011). Initial results from Glasgow are reported in Lawson et al. (2013, forthcoming). Full details of our UTI set-up, the methods for each study, and full analytical results are given in each of the references cited. After brief comments on the technique itself, we show how UTI reveals a probable cause for both the auditory, and the acoustic ambiguities presented by derhoticisation, as well as an articulatory basis for the socially-stratified rhotic-derhotic continuum, in terms of gestural timing (§6.1), tongue configuration (§6.2) and the extent to which these can be accessed (or not) by the listener (§6.3).

5. More may be learnt from a psychoacoustic representation than an acoustic one, given Heselwood & Plug's (2011) recent experiments which strongly suggest that the key perceptual feature of rhoticity (typical of approximants) may be "not a low-frequency F3 *per se*, but rather a single perceptual formant in the F2 region, which we might label F-rho" (p. 870). Lennon's (2011) application of a Bark difference metric (Z3-Z2) to the real-time increase in strong rhotics in middle-class speakers in Glasgow's northern suburbs, suggests that this could be a useful analytical tool for future research.

UTI makes use of ultrasound technology designed for usual medical research, capturing analogue video showing visual dynamic representations of tongue configuration and tongue movement, usually, but not exclusively, in sagittal orientation (see Figure 8).

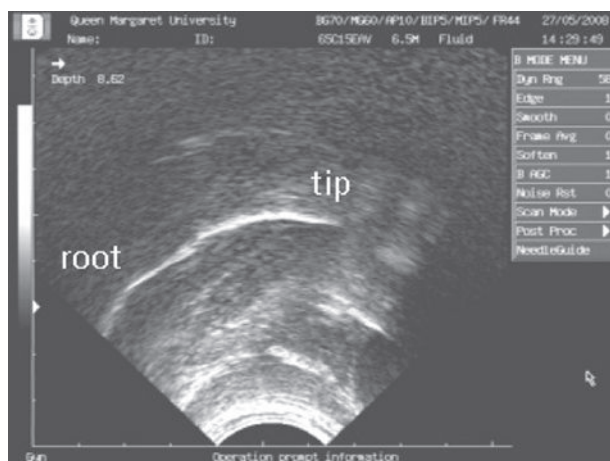


Figure 8. Midsagittal image of the tongue surface produced using a Concept M6 medical ultrasound machine. The tongue root is to the left of the image and the tip is to the right of the image.

In our setup the ultrasound probe is held under the chin by a stabilising headset, and the screen displays a 2D fan-shaped image, showing the water-air interface, i.e. the tongue surface, as a bright white line, thanks to the great intensity of reflections of ultrasound pulses back to the probe. To some extent the internal muscle structure of the tongue can also be seen. It is possible to visualize almost the whole of the mid-sagittal shape and location of the tongue, root, dorsum, front, and sometimes the tongue tip – though the tongue tip is often not visible when it is raised, due to the presence of a sublingual airspace. We use specialist software, *Articulate Assistant Advanced*TM, to capture, process and analyze the data (Articulate Instruments Ltd. 2011).

Whilst UTI gives instant dynamic and static impressions of tongue movement which are immediately informative, quantifying UTI data is challenging and techniques are still under development. Data are also less direct than it might appear, both because of the basic video frame rate (only 30 frames/sec), and the way in which images are constructed by video-output ultrasound machines. This means that ultrasound data are somewhat removed from actual articulation, being both partial and processed. Nevertheless, UTI offers sociophoneticians an excellent

tool for investigating speech articulation, both because it is safe and non-invasive, and because – despite the visible headset and need for technical personnel – the method can have minimal quantitative impacts on speech style. Lawson et al. (2008) shows that in fact stylistic variation is more dependent on the speakers' relationships with their interlocutors, and the presence of friends and peers, than the physical context induced by the equipment. Unlike speakers faced with just a microphone for speech recording, articulatory participants can also be ethically misdirected through a focus on the fact that the recordings are designed to record "changes in the shape of the tongue", which incidentally requires speech.

6.1 Derhoticisation and gestural timing

The UTI data from the pilot data and WL07 corpus uncovered a possible mechanism for derhoticisation in terms of gestural asynchrony. Recall that the auditory transcription was challenging because of the variable percept of sometimes hearing a consonantal gesture and sometimes not, but also from strong pharyngealisation on the vowel.

The articulatory data suggest that derhoticised postvocalic /r/ in our Scottish speech samples involves both (1) an early tongue root retraction gesture and (2) a delayed tongue tip raising gesture, though a systematic study remains to be carried out. An early tongue root retraction gesture could account for the modification of prerhotic vowels, specifically retraction and pharyngealisation of these vowels. The delayed tongue-tip raising gesture means that the maximum of the /r/ gesture is often masked by following consonants, or, prepausally, can occur after the offset of voicing, leaving the /r/ partially or completely inaudible.

This timing, weakening and interarticulatory dissociation of gestures may also account for the weakening of the amplitude of formant energy above F2 observed in the acoustic data. (Exactly how this is achieved is not yet clear, but it seems likely from Stevens' 1998 modeling of the acoustic consequences of the resonating cavities during the production of /r/ and /l/, that the shifts in gestures that we are witnessing are resulting in the formation of an additional cavity with strong damping properties on the spectrum, even before voicing has stopped.) In some speakers, faint dipping of F3 can be seen in a weakly noisy period after voicing has ceased, but this is not always easy to discern and timing of the covert tongue-raising gesture is variable. For example, in Figure 9, the tongue tip only starts to raise in frame 3, just as voicing is ceasing, and then continues to raise during the period of frication; the maximum raising in frame 6 occurs some time after voicing has stopped.

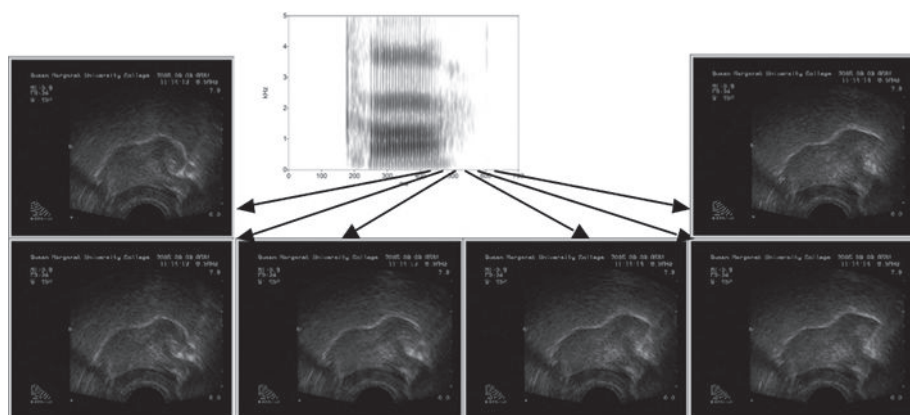


Figure 9. Key UTI frames of an adult male speaker from West Lothian, saying *car* showing a covert tip-raising gesture in the production of coda /r/. The ultrasound images correspond to the time point of the spectrogram. Moving through the frames, it is clear that the tongue front and tip begins to rise after voicing has ceased, and achieves its maximum raising well after.

Thus UTI shows how the timing of two of the gestures contributing to /r/, and in particular their relation to the offset of voicing, means that the primary anterior gesture for the rhotic cannot be reflected in the expected pattern of formant transitions during periodicity. Temporally drifting gestures would also explain the gradient loss of rhoticity. This, and the corresponding shifts in the resonating cavities, help explain the acoustic patterns observed for derhoticizing variants (§5; Figure 6d–f). It is also not surprising that the secondary pharyngealization becomes more audible – the tongue-root gesture is early; and that the /r/ is variably present – the tongue-tip gesture does occur, it just occurs much later, when voicing has stopped.

This account looks at postvocalic /r/ in a particular context, investigated due to previous researchers finding that the phonological environment which most favoured derhoticisation in Scottish English was in stressed, utterance-final position, usually accompanied by vowel breaking, as in e.g. *It's near here* [hiɹ(r)] (see Figure 10, for the distribution of non-rhotic variants; see Romaine 1979: 45; Speitel & Johnson 1983: 28; Lawson et al. 2008).

Figure 10 shows that the second most likely phonological context for derhoticisation was in unstressed syllables, especially in utterance final position, as in Glasgow. Again, this may also relate at least in part to the kind of gestural asynchrony we described above as syllable lengthening is common in utterance-final position, allowing greater gestural asynchrony/dissociation, see Sproat & Fujimura (1993), or possibly also to gestural undershoot, assuming that speakers are likely to be producing an articulated /r/ with a tongue tip gesture, as for

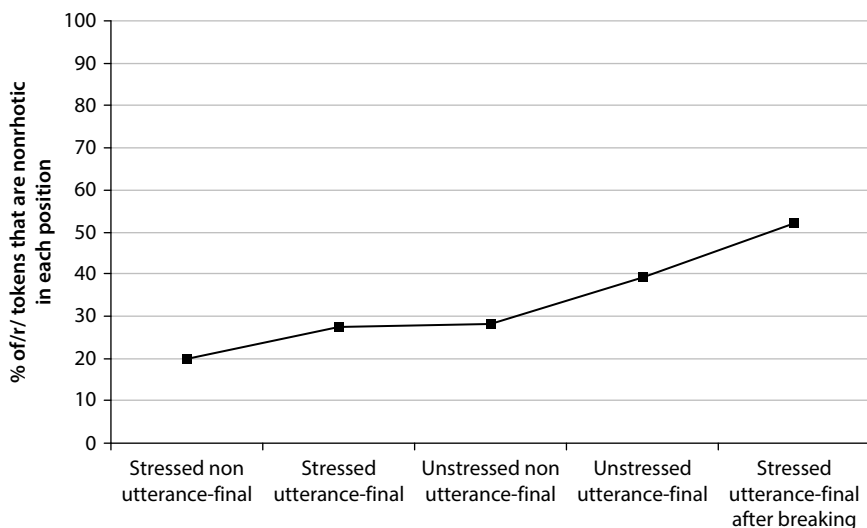


Figure 10. Percentage of (un)stressed tokens in utterance-final and non utterance-final position that were audibly nonrhhotic. $n = 1248$. From Lawson et al. (2008).

e.g. an apical tap. (We have no direct evidence from this particular set of ultrasound data because taps are too fast for the slow frame rate we used). The gradual loss of rhoticity in the history of English English also appears to have started in unstressed syllables (Dobson 1957), and even middle-class speakers who might otherwise be deemed thoroughly rhotic also show audible weakening in this position (Stuart-Smith 2003).

6.2 Tongue configuration and derhoticisation

Derhoticisation probably does not only arise from differences in timing, but also in tongue shape. In the Glasgow 1997 data, those who were likely to derhoticise were also more likely to use taps, if they showed articulated /r/, whereas more rhotic-sounding speakers used more approximants, especially auditorily-strong rhotics which we transcribed as retroflex approximants [ɻ] (Stuart-Smith 2003). Lawson et al. (2011) carried out a further investigation using the eastern Central Belt ECB08 corpus. The design consisted of two parallel analyses of the same data.

The first was an audio-rating analysis of randomized tokens, carried out using the independent classification of tokens via a Praat multiple forced choice interface by two rhotic Scottish-English speakers, both originally from the western Central Belt. Each judge classified the same subset of instances of prepausal postvocalic /r/ (*beer, bear, far, bar, par, purr, fur, for, bore, poor (sure, pure)*), along a 5-point

continuum of auditory ‘strength’ of /r/⁶ (ranging from graded responses such as ‘no /r/’ through ‘derhotic’, ‘alveolar’, ‘retroflex’ to full rhotic vowel, ‘schwar’). The results showed a significant association between the auditory strength of /r/ and social group, and auditory strength of /r/ and gender such that middle-class speakers showed auditorily stronger /r/ than working-class speakers, and girls showed auditorily stronger /r/ than boys (these data are shown in Figure 4 above).

The second – articulatory-rating – analysis of the same data involved the visual classification of the dynamic tongue gestures from the ultrasound videos. Initially, a classification system for tongue-configuration types was devised. This resulted in four categories on a scale from tip-up, through front-up to front bunched and mid-bunched, which takes the differences in configuration for retroflex /r/ and bunched /r/ as effectively lying on a continuum, e.g. Delattre & Freeman (1968) and Zhou et al. (2008). Each video was watched by the second and third authors, and the dynamic configuration of the tongue during the production of each word was noted. Examples of each are shown in the waterfall UTI diagrams in Figure 11. The articulatory-rating study also showed social stratification, with bunched variants occurring mainly in middle-class speech and tip-up variants in working-class speech.

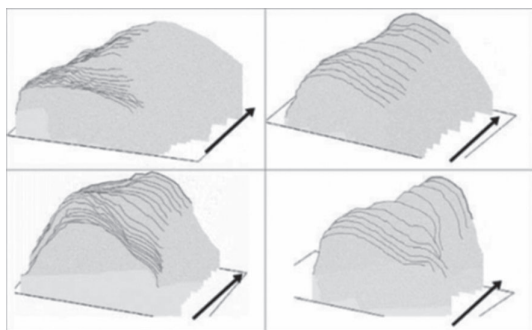


Figure 11. Waterfall diagrams of UTI splines, sampled every 30 ms throughout words ending in /ar/, showing the dynamic movement of the tongue. Time runs in the direction of the arrows. The tongue root is to the left, tongue tip to the right. Top left: tip-up: informant LM16's utterance of *par*; Top right: front-up: LF2's utterance of *far*; Bottom left: front-bunched: EF6's utterance of *far*; Bottom right; mid-bunched: EM5's utterance of *bar*.

6. This was expanded to a 9-point continuum in order to take into account when both raters selected categories that were side by side on the 5-point continuum, i.e. intermediate classification categories were created.

There was a significant correlation ($r=0.637$; $p<0.001$) between the auditory and articulatory ratings. This shows that auditorily weakened /r/, and derhoticisation in the corpus resulted from /r/ articulated with a tongue-tip raised gesture, as discussed above (§6.1), and consistent with the observation in Glasgow of working-class speakers using more taps, and being more derhotic. It also showed that the auditory continuum of rhotic-derhotic has its basis in articulation, since at the other end of the socio-articulatory continuum, the auditorily strongest postvocalic /r/ in these Eastern Central Belt speakers is the result of tongue bunching. Thus both gestural timing and tongue configuration together contribute to the percept of auditorily strong and weak rhotics in the Scottish Central belt.

6.3 Accessing derhoticisation? – Back to the listener

Our articulatory investigations immediately made us wonder how speakers might access, store – and reproduce – such gestures, particularly partially covert tongue-tip gestures when voicing has ceased (§6.1) or the difference between tongue tip-raising and tongue bunching (§6.2). We summarize the results of two relevant small-scale studies below.

Ashton (2011) gauged listener perceptions of articulatorily derhoticised and bunched variants of postvocalic /r/ by investigating whether they were associated with a particular geographical location or socioeconomic status. Auditory stimuli containing postvocalic /r/ were collected from the pre-existing socially-stratified UTI corpora described above and classified according to articulatory gesture (bunched approximant, apical approximant, apical derhoticised /r/ or rless – with no tongue gesture for /r/) and, in the case of derhoticised variants, strength of rhoticity. 16 participants from the Central Belt completed a computer-based subjective reaction test with randomized stimuli. Judgments were made regarding the geographical and social background of the speaker who produced each token. Bunched postvocalic /r/ was found to be strongly associated with middle-class Edinburgh speech, whereas apical approximant /r/ was associated with working-class Scottish speech, but not one particular geographical location. Derhoticised and rless realisations of postvocalic /r/ were found to be associated with Glasgow, and derhoticisation was strongly associated with working-class speech.

Lawson et al. (2011b) presents preliminary evidence for configurational lingual adaptation in Scottish postvocalic /r/ during mimicry. A male speaker, originally from the west of Scotland, was asked to mimic a number of audio stimuli extracted from the ECB08 and WL07 corpora. The articulatory gestures underlying the audio stimuli were known. His mimicked articulations were then compared to his baseline UTI recordings (only a small number of items could be

compared). The mimicked data showed little adaptation of tongue configuration, but some shift in the timing of the gesture (with respect to offset of voicing) particularly when responding to the tongue-bunched auditory stimuli. It was also interesting to note that the covert, delayed apical /r/ gesture was not reproduced when mimicking the audio signal from the derhoticised utterance of *hurt*; instead the speaker produced an rless word *hut* (see Figure 12).



Figure 12. Waterfall diagrams of UTI splines from the mimicking study (Lawson et al. 2011b). Left: the original production of *hurt* by the mimicker, which sounds weakly rhotic. Middle: the production of the stimulus for mimicking, auditorily derhoticised *hurt*, but with covert delayed tongue-tip raising. Right: the mimicked production of *hurt*, without any tongue-tip raising, and sounding like *hut*. (Note that /t/ in these word is realized as a glottal stop.)

With respect to derhoticisation, the results confirmed that delay in the tongue tip gesture can lead to an ambiguous auditory percept not only for an analyst, but also for a derhoticising member of the Scottish speech community. Our suspicions that the acoustic signal could be difficult to parse seem plausible, though this needs more investigation, which is now underway in a systematic socio-articulatory phonetic study using mimicking in conjunction with UTI recordings. This study will allow us to develop a clearer picture of how articulatory variation spreads from speaker to hearer.

7. Discussion and reflection: The sociophonology of Scottish derhoticisation

The studies presented, both by previous scholars and ourselves, show that rhoticity in Scottish English has been eroding gradually over the 20th century for working-class speakers, and possibly for longer. This is counterbalanced by an increasingly auditorily strong rhoticity in middle-class speakers (see Lennon 2011). The changes are largely driven by sociolinguistic dynamics within this Scottish community, though there is evidence for reinforcement from an unlikely source, indirect contact with London English on TV. Describing and accounting for these changes phonetically has also been a focus – and of course – is far from complete. A number of issues arise, but we focus here on two which relate to representation,

the first concentrating on analysts and how we are to deal with such labile data, the second, making suggestions about the possible mental representations held by speakers and hearers participating in these changes.

7.1 Analytical representation of sociophonetic variation: The speaker-hearer triangle

We illustrate some of the implications of our articulatory investigation for the analytical representation of variation by focusing on the middle-class, rhotic, end of continuum.

The audio-/articulatory-rating study shows clearly that auditory judgments result in auditory objects, and not the quasi-articulatory objects suggested by IPA representations. Recall that auditorily-strong approximants in middle-class speakers were consistently phonetically transcribed as ‘retroflex’ using the IPA symbol [ɻ] (e.g. Johnston 1997). However, the UTI data show that the actual configuration for these variants is likely to involve tongue-bunching, with no tip raising at all.

It is also clear that – at some level at least – the differences between tongue-tip raising and tongue bunching can be discerned by members of this speech community, since they show systematic patterning with social membership of particular subgroups. This shows that the fine-grained differences in /r/ production can be exploited and used to construct and reflect social meaning (Eckert 2008). Being an urban middle-class girl involves the use of a specific kind of auditorily-strong, bunched /r/; at the opposite pole, working-class girls in the western Central Belt are continuing to use non-rhotic and derhoticised variants. It is clear that the phonological category of /r/ in this position is closely linked with locally-situated social categories.

Moreover, these results for Scottish English are in contrast to those found for American English /r/ by Twist et al. (2007), where listeners were found to be ‘at best weakly aware’ of articulatory variation (retroflexion and bunching) in /r/ (Twist et al. 2007: 215). However, there are good reasons to assume that bunched and retroflex /r/ could be perceptually distinguished. Johnson (2011) points out that Zhou et al. (2008) identify clear acoustic differences in the frequency and trajectory of F3 and F4 between the two variants. He demonstrates their perceptual salience by showing that acoustic stimuli created with these differences can lead to differential perceptual compensation. In addition, even if acoustic equivalence is assumed for different articulatory strategies, the coarticulatory effect of these very different /r/ articulations may provide the listener with information regarding differences in underlying articulation, see Lawson et al. (2013). This suggests that rather fine-grained phonetic differences (the higher formants may often be only

weakly resonated), which are potentially accessible, have been exploited for social meaning in Scottish English, but seem to remain unattached in American English.

More generally, using these different phonetic representations of postvocalic /r/ provides a good illustration of what we call here the ‘speaker-hearer triangle’, composed of auditory, acoustic, and articulatory representations (Ogden 2009; Heselwood & Plug 2011 looks at auditory, acoustic and psychoacoustic views of rhoticity). Figure 13 shows auditorily strong postvocalic /r/: each representation gives a different picture of the ‘same’ phenomenon. In many ways each is as valid as the other, and of course, as we have seen, they are all interconnected but not necessarily in straightforward ways.

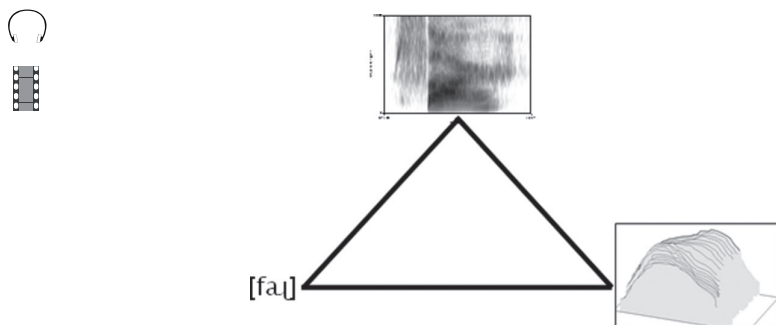


Figure 13. An illustration of the ‘speaker-hearer triangle’ of auditory, acoustic, and articulatory representations of the auditorily-strong postvocalic /r/ in a middle-class Edinburgh girl’s production of the word *far*.

Ideally, representing sociophonetic variation would be able to refer to all three dimensions of the speaker-hearer triangle. Adding articulatory data can prove very fruitful (also Wright & Kerswill’s 1989 conclusions for using electropalatographic data in conjunction with auditory transcription). This can also help us to reflect on the different kinds of representation – and their intersections – that might be involved in the transmission and propagation of sociophonetic variation, which is also of crucial importance in modeling language variation and change (e.g. Marotta this volume). The traditional notion of the speaker-hearer chain (e.g. Denes & Pinson 1993) assumes that articulatory gestures from the speaker give rise to acoustic objects, which in turn become auditory objects for the listener to decode (see also Ohala, e.g. 1989). How variation which appears to be so auditorily subtle, yet can be acquired and transmitted such that it can carry social meaning for a community, requires substantial further investigation.

7.2 Mental representation of sociophonetic variation: A symbolic relationship?

It is clear that the rhotic-derhotic continuum in Scottish English in the Central Belt is undergoing shifts in fine phonetic realization. It is also clear that it is impossible to describe the scope of phonological rhoticity without reference to social factors, both macro and micro. For these speakers /r/ in this position is not just an /r/, it is always a certain kind of socially-embedded /r/; at the descriptive level it is extremely difficult to separate the phonological from the social. It is also difficult to assume that these entities do not relate to each other very closely for speaker-hearers. Data of this kind demand phonological representations which recognize the interconnected relationships between social and phonological variation which speakers in these communities need to store, control, access, and acquire.

The approach which recognizes such connections and which ‘embeds indexicality centrally within phonological knowledge’ (Foulkes & Docherty 2006: 426), is the range of theories of phonological representation grouped under the term ‘Exemplar Theory’ (e.g. Goldinger 1998; Johnson 2006; Hawkins 2003). These models share the assumption that phonological representations are based in some way on stored experiences of speech (‘exemplars’), memory clouds across which abstractions are probabilistically derived. Increasing emphasis is placed on the need for abstractions accrued from exemplar memory (corresponding to phonological categories in other perception-production models) being stored concurrently and with connections to exemplars, so-called ‘hybrid’ models (Goldinger 2007; Pierrehumbert 2006).

The results from the rhotic-derhotic continuum in Scottish English also have implications for hybrid models, particularly with respect to the relationships between phonological and social detail and abstraction. Schematic accounts of exemplar-based representations such as that by Johnson (2006) distinguish the exemplar map from accruing abstractions, but interestingly also make a separation between phonological and social categories at the abstract level. This implies that the connections between these two kinds of abstractions (as well as with others) are always made through exemplar memory. But it is clear that phonological abstractions such as ‘postvocalic /r/’, which are accessible to speakers especially through stereotypes, also relate to social abstractions at the same time. Moreover if we consider the acquisition of speech variation which is necessarily socially-embedded (e.g. Foulkes et al. 2005; Labov this volume), it seems difficult to assume that the emerging abstractions are not linked – or linkable – if only because the shared/simultaneous activation of phonological and social categories would be so frequent. Rather these sociophonetic data, and those from many other

sociophonetic studies (e.g. Foulkes et al. 2010), suggest that these abstract levels likely relate to each other directly, as the result of persistent coupling in the system.

If we make this assumption (and it seems inevitable that we must),⁷ an analogy from ancient Greek society may be useful for considering the possible nature of the relationship between these two abstract levels. Greek *symbola* were originally two halves of the same object, each a *symbolon*, which could be fitted together for purposes of personal recognition (Herman 1987). Only later in the classical period did the meaning of the word *symbolon* shift from denoting part of a two-part tally, to tokens which could be used like tickets in exchange for goods, continued in English 'symbolic'. The original *symbolon/symbola* relationship had two key aspects: (1) each *symbolon* could and did exist separately, for example, members of a dispersed family could keep them for a long time; but each *symbolon* was only meaningful when reunited with its partner (*symbola*). (2) *Symbola* could be formally similar, but each half could also be different from each other (Harris 2000: 23).

The relationship between phonological and social abstractions emerging from exemplar memory could be likened to the *symbolon/symbola* relationship.⁸ Both kinds of categorization, at whatever level, can and do exist separately, both for analysts, and for speakers under particular conditions. For example, it is clearly possible to undertake separate analyses of phonological structure, or of social categorization, without reference to each other (Labov 2006). Speakers too can access phonological categories without reference to social categories, e.g. in psycholinguistic manipulation tasks, and social stereotypes can be retrieved without automatically referring to speech. But we suggest that the usual situation for speakers in daily interaction is that the social and phonological systems function in a *symbola* relationship, namely they are linked or continually linking such that "each is significant ... as a counterpart of the other" (Harris 2000: 23). Such an analogy allows us to think about the social and phonological systems as having a separate, yet co-emergent relationship at the abstract level. The links themselves would be established through co-ordinated simultaneous activation, leading to persistent coupling within and across the exemplar map, and hence the entrenchment of linked/linkable social and phonological categories (this kind of modeling assumes activation and resonance discussed by Johnson 2006). At the same time, prior knowledge encapsulated in such social-phonological linkages will serve to mediate the treatment of subsequent input exemplars (Goldinger 2007).

7. Keith Johnson (p.c. 2011) notes the difficulty of two-dimensional graphical representations of phenomena and processes that are (a) multidimensional and (b) thoroughly inter-related.

8. The *symbola* relationship could also be used metaphorically to refer to the special relationships between entities; Aristotle's account of speech and writing is given in these terms in the introduction to his *De Interpretatione*, 16^a3–8.

8. Conclusions

This paper has taken an aspect of Scottish English phonology, postvocalic /r/, which appears to be changing. A strictly phonetic account of this phenomenon is not possible; social information is required even to be able to specify the kinds of phonetic variation which are emerging. Phonetic analysis carried out on socially-stratified speech data from the Scottish Central Belt shows how speakers at different ends of the social spectrum are exploiting very fine phonetic differences in coda /r/ for social ends. Conversely, such socially-informed phonetic analyses sheds light on the mechanisms of the weakening and derhoticisation of /r/, and its auditory strengthening. We discuss representations of speech from three points of the possible ‘speaker-hearer triangle’: all three give partial impressions of the phenomena. All three are needed together in order to gain an improved understanding of their nature.

Variation and change in coda /r/ is also informative for sociolinguistic theory. Unlike middle-class non-rhoticity, working-class derhoticisation in Scottish English has never been interpreted as a contact-induced change through interacting with non-rhotic English English speakers. But our results show that strong psychological engagement with a London-based television show is linked to increased r-lessness. This strongly suggests that current models of media influence, which assume that the prior knowledge of the viewer is essential, should also be extended to language, and specifically that prior sociolinguistic knowledge of the viewer may act as a sociolinguistic filter on incoming media language – leading to decay or enhancement depending on the degree of social relevance and linguistic congruence with the speaker/viewer’s system.

Overall, we have learnt a lot, but there is still much more to discover, both about this particular phenomenon, and about some of the wider issues which it exemplifies, for example:

- What has happened in real time over the past century? Are we witnessing language change, and if so how fast or gradual is this? Only empirical study of real-time data can begin to answer this question.
- How can we objectively describe and assess derhoticisation? We need improved understanding of the acoustics, and the psychoacoustics, of rhoticity.
- How do changes in coda position relate to those in onset position? Our study focuses on coda /r/, particularly in utterance-final position. We have noted that this location seems to be particularly salient socially. More work needs to be carried out – like that of Pukli and Jauriberry (2011) – which analyses /r/ in all positions.

- How does subtle articulatory variation of this kind get transmitted? Modeling mechanisms of language variation and change rely on a much improved understanding of the relationships between speakers and hearers, and in particular, how hearers may respond to input from speakers at the level of articulation.
- How do speakers phonetically and socially decode speech experienced without the possibility of interaction? This area is virtually unresearched, but needs to be explored empirically if we are to make progress in understanding how engaging with the broadcast media relates to spoken language in the community.

Our current and future research, with each other and other colleagues, aims to try to tackle some of these questions. But it is now clear to us, after working on this phenomenon for over 15 years, that what appears to be the answer is usually the starting point for more questions: in fact this particular sociophonetic journey has only just begun.

References

- Aitken, Adam J. 1984. "Scots and English in Scotland". *Language in the British Isles* ed. by Peter Trudgill, 517–532. Cambridge: Cambridge University Press.
- Anthony, A., D. Bogle, T. T. S. Ingram & M. W. McIsaac. 1971. *The Edinburgh Articulation Test*. Edinburgh: Churchill Livingstone.
- Articulate Instruments Ltd. 2011. *Articulate Assistant Advanced User Guide: Version 2.13*. Edinburgh: Articulate Instruments Ltd.
- Ashton, Lindsay. 2011. *Perception of /r/ in Scottish English*. Undergraduate Dissertation, Edinburgh, Queen Margaret University.
- Bucholtz, Mary. 2009. "From stance to style: gender, interaction, and indexicality in Mexican immigrant youth slang". *Stance. Sociolinguistic Perspectives on Stance* ed. by Alexandra Jaffe, 146–170. New York: Oxford University Press.
- Carey, Emma. 2010. *Cross Accent Perception of Standard Southern British English and Glasgow English*, Undergraduate Dissertation, University of Glasgow.
- Coupland, Nikolas. 2007. *Style: Language variation and identity*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511755064
- Denes, Peter B. & Elliot N. Pinson. 1993. *The Speech Chain: The Physics and Biology of Spoken Language*. Oxford: W. H. Freeman and Company.
- Delattre, Pierre & Donald C. Freeman. 1968. "A dialect study of American r's by x-ray motion picture". *Linguistics* 44. 29–68.
- Dobson, Eric J. 1957. *English Pronunciation 1500–1700*. Oxford: Clarendon Press.
- Docherty, Gerard J. & Paul Foulkes. 1999. "Derby and Newcastle: instrumental phonetics and variationist studies". *Urban voices: variation and change in British accents* ed. by Paul Foulkes & Gerard J. Docherty, 47–71. London: Arnold.

- Douglas, Fiona M. 2009. *Scottish Newspapers, Language and Identity*. Edinburgh: Edinburgh University Press. DOI: 10.3366/edinburgh/9780748624379.001.0001
- Durand, Jacques. 2004. "English in early 21st century Scotland: a phonological perspective". *La Tribune Internationale des Langues Vivantes* 36. 87–108.
- Eckert, Penelope. 2000. *Linguistic Variation as Social Practice*. Oxford: Blackwell.
- Eckert, Penelope. 2008. "Variation and the indexical field". *Journal of Sociolinguistics* 12(4). 453–76. DOI: 10.1111/j.1467-9841.2008.00374.x
- Foulkes, Paul and Gerard J. Docherty. 2006. "The social life of phonetics and phonology". *Journal of Phonetics* 34. 409–38. DOI: 10.1016/j.wocn.2005.08.002
- Foulkes, Paul, Gerard J. Docherty & Dominic J. L. Watt. 2005. "Phonological variation in child directed speech". *Language* 81. 177–206. DOI: 10.1353/lan.2005.0018
- Foulkes, Paul, James M. Scobbie & Dominic J. L. Watt. 2010. "Sociophonetics". *Handbook of Phonetic Sciences* (2nd ed.) ed. by William Hardcastle, John Laver & Fiona Gibbon, 703–754. Oxford: Blackwell. DOI: 10.1002/9781444317251.ch19
- Goldinger, Stephen D. 1998. "Echoes of echoes? An episodic theory of lexical access". *Psychological Review* 105. 251–79. DOI: 10.1037/0033-295X.105.2.251
- Goldinger, Stephen D. 2007. "A complementary-systems approach to abstract and episodic speech perception". *Proceedings of the 16th International Congress of Phonetic Sciences* ed. by Jürgen Trouvain & William J. Barry, 49–54 (ID 1781). <http://www.icphs2007.de/>
- Grant, William. 1914. *The Pronunciation of English in Scotland*. Cambridge: Cambridge University Press.
- Grant, William and James M. Dixon. 1921. *A Manual of Modern Scots*. Cambridge: Cambridge University Press.
- Gunter, Barrie. 2000. *Media research methods: Measuring audiences, reactions and impact*. London: Sage.
- Hall, Stuart. 1980. "Encoding/decoding". *Culture, Media, Language: Working Papers on Cultural Studies, 1972–79*, 128–38. London: Hutchinson.
- Harris, Roy. 2000. *Rethinking writing*. London: Athlone Press.
- Hawkins, Sarah. 2003. "Roles and representations of systematic fine detail in speech understanding". *Journal of Phonetics* 31. 373–405. DOI: 10.1016/j.wocn.2003.09.006
- Herman, Gabriel. 1987. *Ritualised Friendship and the Greek City*. Cambridge: Cambridge University Press.
- Heselwood, Barry and Leendert Plug. 2011. "The role of F2 and F3 in the perception of rhoticity: Evidence from listening experiments". *Proceedings of the XVIII International Congress of Phonetic Sciences*, 867–70. http://www.icphs2011.hk/ICPHS_CongressProceedings.htm
- Jaffe, Alexandra. 2009. "Introduction: The sociolinguistics of stance". *Stance. Sociolinguistic Perspectives on Stance* ed. by Alexandra Jaffe. New York: Oxford University Press.
- Johnson, Keith. 2006. "Resonance in an exemplar lexicon: The emergence of social identity and phonology". *Journal of Phonetics* 34. 485–499. DOI: 10.1016/j.wocn.2005.08.004
- Johnson, Keith. 2011. 'Retroflex versus bunched /r/ in compensation for coarticulation', *UC Berkeley Phonology Lab Annual Report*. 114–27.
- Johnston, Paul. 1985. "The rise and fall of the Morningside/Kelvinside accent". *Focus on Scotland* ed. by Manfred Gorch, 37–56. Amsterdam: Benjamin.
- Johnston, Paul. 1997. "Regional Variation". *The Edinburgh History of Scots* ed. by Charles Jones, 433–513. Edinburgh: Edinburgh University Press.
- Kuhl, Patricia. 2010. "Brain mechanisms in early language acquisition". *Neuron* 67. 713–775. DOI: 10.1016/j.neuron.2010.08.038

- Labov, William. 1972. *Sociolinguistic patterns*. Oxford: Blackwell.
- Labov, William. 1994. *Principles of Linguistic Change, vol. I, Internal Factors*. Oxford: Blackwell.
- Lawson, Eleanor, Jane Stuart-Smith & James M. Scobbie. 2008. "Articulatory insights into language variation and change: Preliminary findings from an ultrasound study of derhoticisation in Scottish English". *University of Pennsylvania Working Papers in Linguistics* 14(2). 101–110.
- Lawson, Eleanor, James M. Scobbie & Jane Stuart-Smith. 2011. "The social stratification of tongue shape for postvocalic /r/ in Scottish English". *Journal of Sociolinguistics* 15(2). 256–268. DOI: 10.1111/j.1467-9841.2011.00464.x
- Lawson, Eleanor, Jane Stuart-Smith, James M. Scobbie, Malcah Yaeger-Dror & Margaret MacLagan. 2011a. "Liquids". *Sociophonetics: A Student's Guide* ed. by Malcah Yaeger-Dror & Marianna di Paolo, 72–86. London: Routledge.
- Lawson, Eleanor, James M. Scobbie & Jane Stuart-Smith. 2011b. "A single case study of articulatory adaptation during acoustic mimicry". *Proceedings of the XVIIth International Congress of Phonetic Sciences*, 1170–1173. http://www.icphs2011.hk/ICPHS_CongressProceedings.htm
- Lawson, Eleanor, James M. Scobbie & Jane Stuart-Smith. 2013. "Bunched /r/ promotes vowel merger to schwar: An Ultrasound Tongue Imaging study of Scottish sociophonetic variation." *Journal of Phonetics* 41. 298–210. DOI: 10.1016/j.wocn.2013.01.004
- Lawson, Eleanor, James M. Scobbie & Jane and Stuart-Smith. 2013, forthcoming. "A socio-articulatory study of Scottish rhoticity". *Sociolinguistics in Scotland* ed. by Robert Lawson, Basingstoke: Palgrave.
- Lennon, Robert. 2011. *A real-time study of rhoticity in Glaswegian between 1997 and 2011*. Undergraduate Dissertation, University of Glasgow.
- Llamas, Carmen. 2010. "Convergence and divergence across a national border". *Language and Identities* ed. by Carmen Llamas & Dominic Watt, 227–236. Edinburgh: Edinburgh University Press.
- Llamas, Carmen, Dominic Watt & Daniel Ezra Johnson. 2009. "Linguistic Accommodation and the Salience of National Identity Markers in a Border Town". *Language and Social Psychology* 28. 381–407. DOI: 10.1177/0261927X09341962
- Macafee, Caroline. 1983. *Varieties of English around the World: Glasgow*. Amsterdam: John Benjamins.
- McAllister, Anne H. 1938. *A Year's Course in Speech Training*. London: Hodder and Stoughton.
- MacFarlane, Andrew E. & Jane Stuart-Smith. 2012. "'One of them sounds sort of Glasgow Uni-ish': Social judgements and fine phonetic variation in Glasgow". *Lingua* 122(7). 764–778. DOI: 10.1016/j.lingua.2012.01.007
- Ohala, John J. 1989. "Sound change is drawn from a pool of synchronic variation". L. E. Breivik & E. H. Jahr (eds.), *Language Change: Contributions to the study of its causes* ed. by Leiv E. Breivik & Ernst H. Jahr, 173–198. Berlin: Mouton de Gruyter.
- Pierrehumbert, Janet. 2006. "The next toolkit". *Journal of Phonetics* 34. 516–530. DOI: 10.1016/j.wocn.2006.06.003
- Plug, Leendert & Richard Ogden. 2003. "A parametric approach to the phonetics of postvocalic /r/ in Dutch". *Phonetica* 60. 159–86. DOI: 10.1159/000073501
- Pukli, Monika and Thomas Jauriberry. 2011. "Language change in action – variation in Scottish English". *Recherches Anglaises et Nord-Américaines* 44. 83–100.
- Richmond, Suzanne. 2013. *A Real-Time Sociophonetic Study of Scottish English: The Realisation of /r/ across a Century*. Unpublished Honours Dissertation, University of Glasgow.

- Romaine, Suzanne. 1978. "Postvocalic /r/ in Scottish English: sound change in progress?". *Sociolinguistic patterns in British English* ed. by Peter Trudgill, 144–57. London: Edward Arnold.
- Scobbie, James M. 2007. "Biological and social grounding of phonology: Variation as a research tool". *Proceedings of the 16th International Congress of Phonetic Sciences* ed. by Jürgen Trouvain & William J. Barry, 211–214 (ID 1765). <http://www.icphs2007.de/>
- Scobbie, James M., Olga B. Gordeeva & Benjamin Matthews. 2007. "Scottish English Speech Acquisition". *The International Guide to Speech Acquisition* ed. by Sharynne McLeod, 221–240. Thomson Delmar Publishing.
- Scobbie, James M. and Koen Sebrechts. 2011. "Acoustic, articulatory and phonological perspectives on rhoticity and /r/ in Dutch (Eds.) *Interfaces in Linguistics: New Research Perspectives* ed. by Raffaella Folli & Christiane Ulbrich, 257–277. Oxford: Oxford University Press.
- Speitel, Hans-Henning & Paul Johnston. 1983. "A Sociolinguistic investigation of Edinburgh speech". Unpublished final report (Grant No. 000230023) for the Social Science Research Council (Great Britain).
- Sproat, Richard & Osamu Fujimura. 1993. "Allophonic variation in English /l/ and its implications for phonetic implementation". *Journal of Phonetics* 21. 291–311.
- Stevens, Kenneth. 1998. *Acoustic Phonetics*. Cambridge, MA: MIT Press.
- Stuart-Smith, Jane. 1999. "Glasgow: accent and voice quality". *Urban voices: variation and change in British accents* ed. by Paul Foulkes & Gerard Docherty, 201–222. London: Arnold.
- Stuart-Smith, Jane. 2003. "The phonology of Modern Urban Scots". *The Edinburgh Companion to Scots* ed. by John Corbett, Derrick McClure & Jane Stuart-Smith, 110–137. Edinburgh: Edinburgh University Press.
- Stuart-Smith, Jane. 2006. "The influence of media on language". *The Routledge Companion to Sociolinguistics* ed. by Carmen Llamas, Louise Mullany & Peter Stockwell, 140–148. London: Routledge.
- Stuart-Smith, Jane. 2007. "A sociophonetic investigation of postvocalic /r/ in Glaswegian adolescents". *Proceedings of the 16th International Congress of Phonetic Sciences* ed. by Jürgen Trouvain & William J. Barry, 211–214 (ID 1307). <http://www.icphs2007.de/>
- Stuart-Smith, Jane. 2011. "The view from the couch: changing perspectives on the role of the television in changing language ideologies and use". *Standard Languages and Language Standards in a Changing Europe* ed. by Tore Kristiansen & Nikolas Coupland, 223–239. Oslo: Novus.
- Stuart-Smith, Jane & Claire Timmins. 2010. "The role of the individual in language change". *Language and Identity* ed. by Carmen Llamas & Dominic Watt, 39–54. Edinburgh: Edinburgh University Press.
- Stuart-Smith, Jane, Claire Timmins & Fiona Tweedie. 2007. "Talkin' Jockney: Accent change in Glaswegian". *Journal of Sociolinguistics* 11. 221–61.
- Stuart-Smith, Jane, Claire Timmins, Gwilym Pryce & Barrie Gunter. 2013. "Television is also a factor in language change: Evidence from an urban dialect". *Language* 89 (3). 1–36.
- Trudgill, Peter. 1986. *Dialects in contact*. Oxford: Blackwell.
- Twist, Alina, Adam Baker, Jeff Mielke & Diana Archangeli. 2007. "Are 'covert' /r/ allophones really indistinguishable?". *University of Pennsylvania Working Papers in Linguistics* 13(2). 207–216. <http://repository.upenn.edu/pwpl/vol13/iss2/16>
- Wells, John C. 1982. *Accents of English*. Cambridge, Cambridge University Press.
- Williams, Irene F. 1909. *Phonetics for Scottish students: the sounds of polite Scottish described and compared with those of polite English*. Glasgow: J. MacLehose.

- Wright, Susan & Paul Kerswill. 1989. "Electropalatography in the study of connected speech processes". *Clinical Linguistics and Phonetics* 3. 49–57. DOI: 10.3109/02699208908985270
- Yaeger-Dror, Malcah, Tyler Kendall, Paul Foulkes, Dominic Watt, Jillian Oddie, Phil Harrison & Colleen Kavenagh. 2009. "Perception of r-fulness by trained listeners". Paper presented at NWAV 37, Houston, TX.
- Zhou, Xinhui, Carol Espy-Wilson, Mark Tiede, Suzanne Boyce, Christy Holland & Ann Choe. 2008. "A magnetic resonance imaging-based articulatory and acoustic study of 'retroflex' and 'bunched' American English /r/". *Journal of the Acoustical Society of America* 123(6). 4466–4481. DOI: 10.1121/1.2902168

Where and what is (t,d)?

A case study in taking a step back in order to advance sociophonetics

Rosalind A.M. Temple
New College, Oxford

The variable deletion of /t,d/ in word-final clusters in English has garnered much attention from sociolinguists and, more recently, phonologists, most of whom model it as a binary variable phonological rule. This paper examines in detail some (t,d) clusters in York English and compares them with other word-final singleton and cluster consonants. In the light of the general literature on English, it explores an alternative view, that in at least one variety of British English “-t,d deletion” is in fact a function of the common connected speech processes which apply at the boundaries between words. It thus underlines the importance for advances in sociophonetics of taking a step back to examine critically the basic units of analysis of variable rules.

1. Introduction¹

The ultimate aim of sociophonetics, consistent with the vision of variationist sociolinguistics more broadly since its inception, goes beyond mapping the distribution of variants across social categories to integrating variability into the grammar (see, for example, Stuart-Smith et al.’s discussion of exemplar theory in this volume). At the heart of sociophonetics is phonetic detail, and the crucial

1. My heartfelt thanks to the following colleagues and friends for encouraging, advising and challenging me during the preparation of this paper: at Pisa, Gillian Sankoff and Jane Stuart-Smith; in Oxford, John Coleman and Sali Tagliamonte; remotely and elsewhere, Ricardo Bermúdez-Otero and John Glyn. Thanks too to the editors and reviewers of this volume for their painstaking reading of the first draft of the paper. Its shortcomings remain, of course, my own.

contribution of the field, if it may be called such,² is the accurate description of patterns of phonetic variation in naturalistic data, on which theoretical constructs may be built. However, notwithstanding the work of Stuart-Smith, Scobbie and their collaborators, reported here and elsewhere, and phonetically informed variationist analyses of, for example, *t*-glottalisation (e.g. Docherty & Foulkes 2005), insufficient attention has been paid to the phonetic substance of some major consonantal variables. The present paper focuses on one such variable, perhaps the most widely studied consonantal variable in English sociolinguistics, whose social indexicality has been shown to be restricted to relatively few dialects, but which has garnered so much attention within and beyond sociolinguistics because of its claimed implications for phonological theory.

The variable deletion of coronal stops in word-final clusters (e.g. *stopped* pronounced as variably [stɒp̚t̚] or [stɒp̚]) seems to occur in all varieties of English and has been one of the most studied variables in the variationist sociolinguistics of the language. It has been used as a diagnostic in debates about the origins of African American Vernacular English (AAVE) since the late 1960s (e.g. Wolfram 1969) and more recently it has figured prominently in the exploration of cross-dialectal differences (e.g. Santa Ana 1992; Smith et al. 2009), the acquisition of variable constraints (e.g. Guy & Boyd 1990; Roberts 1997; Smith et al. 2009) and particularly the relationship between variation and phonological theory (e.g. Guy 1991; Guy & Boberg 1997; Bermúdez-Otero 2010a and 2010b; Coetzee & Pater 2011). The phonological model most widely applied to the variable has been one rooted in Lexical Phonology (LP), which characterises (t,d)³ as an iterative derivational rule that applies variably in the lexical and postlexical phonology. The analysis is motivated crucially by there being a consistent (statistical) morphological constraint on (t,d) whereby monomorphemic forms undergo deletion of the final consonant considerably more frequently than bimorphemic forms. However, findings from several recent studies (e.g. Tagliamonte & Temple 2005; Smith et al. 2009; Guy et al. 2008; Hazen 2011) have introduced an element of doubt as to the role of this particular constraint, thus undermining the LP account of the variable. Temple (ms) goes a step further in an exploration of some of the theoretical and methodological issues which arose during the research reported in Tagliamonte & Temple (2005), arguing that once the morphological constraint is called into

2. See Celata & Calamai and Stuart-Smith et al., this volume, for brief discussions of the scope of the term 'sociophonetics'.

3. The variable notation will be used here as a shorthand means of referring to both the variable rule which deletes word-final coronal stops in clusters and the set of consonants affected by that rule.

question the case for treating (t,d) as a phonologically categorical⁴ variable rule within any framework needs to be made anew, since there remain no obvious grounds for treating it in this way.⁵ Moreover, the phonetic issues highlighted in that paper suggest that there are good grounds for treating (t,d) as a function of common Connected Speech Processes (CSPs) observed by many phoneticians in English rather than a particular variable rule restricted to these coronal clusters.

The present paper will attempt to make the case for the CSP view of (t,d) through a qualitative re-examination of data from some of the 38 speakers analysed by Tagliamonte & Temple (2005), together with comparable data from the same corpus containing other underlying coda consonant clusters and singleton consonants. The data are all taken from audio recordings of sociolinguistic interviews collected for the York Corpus of British English under the direction of Sali Tagliamonte and described in Tagliamonte (1998).⁶ As Stuart-Smith et al., this volume, demonstrate in their analysis of the complex social indexicality of the detailed phonetics of rhotics, even cutting edge articulatory techniques cannot in isolation give us a full picture of sociophonetic variability and need to be triangulated with auditory and acoustic analyses, which are themselves imperfect representations. Articulatory data are not available for the York recordings, so the analysis in this paper will draw on acoustic and auditory observations, illustrated by detailed phonetic transcriptions and a small sample of illustrative spectrograms; however, since the issues raised also crucially concern articulations which are not necessarily audible or observable from the acoustic signal, reference will be made throughout to the literature reporting relevant articulatory studies.

4. There is a mismatch between the use of the word 'categorical' by variationists on the one hand and general phonologists on the other: the former oppose 'categorical' rules, which always apply (as in cases of regular allophony) to 'variable rules', which apply probabilistically (e.g. coronal stop deletion is more likely to occur before consonants than vowels); the latter differentiate between 'categorical' processes (e.g. the 'replacement' of a voiced stop by a voiceless one under assimilation) and 'gradient' ones (e.g. the partial devoicing to various degrees of a voiced stop under the same conditions). Both dichotomies apply to the discussion of (t,d) but the term 'categorical' is used here to mean non-gradient, since all analyses of the variable in question agree that it is probabilistic.

5. Some scholars (e.g. Bermúdez-Otero 2010a; Myers 1996) argue for a dual view of (t,d) as both a categorical and a gradient rule, as explicitly allowed for in Kiparsky's (1985) view of LP. The positive case for the categorical rule still needs to be made under this view.

6. The data collection was funded by a research grant (#R000238287) from the Economic and Social Research Council for the United Kingdom. Digitisation of a subset of the data for the present paper was funded by the John Fell fund of the University of Oxford. I am grateful to Damien Mooney for his efficient assistance with the digitisation.

In §2, I examine a range of CSPs to ascertain whether the range of phonetic patterns found in (t,d) consonants is consistent with a CSP analysis and whether these patterns are exclusive to (t,d) consonants. The analysis will touch on issues which must be taken into account in deciding whether word-final clusters and/or other CSPs are amenable to analysis in terms of variable rules. These are issues which have long been the subject of discussion in the phonetics literature and they have not gone entirely unnoticed in discussions of (t,d), having been raised by e.g. Wolfram (1993), but there is little subsequent evidence that Wolfram's concerns have been heeded. In the discussion in §3, I turn to the implications of these observations for modelling the behaviour of word-final stop consonants in the grammar in the light of ongoing debates about the phonetics-phonology interface, a prerequisite to sociophonetic/sociophonological modelling. I thus hope to demonstrate how, paradoxically, advances in sociophonetics might sometimes be achieved by stepping back and re-examining the phonetic detail behind a rule which is generally held to be predominantly a categorical phonological one. It will be seen that much can emerge from such an apparently retrospective approach which can contribute to advances in sociophonetics and wider debates concerning the relationship of its findings to phonetic and phonological theory, albeit there are questions which will remain unanswered until further advances are made by applying particularly articulatory techniques to this variable.

2. (t,d) and Connected Speech Processes⁷

In contrast to the phonologically based accounts of (t,d), which posit a categorical alternation between the presence and absence of a surface reflex of underlying word-final /t,d/, CSPs provide, in Nolan's words, "a way of describing a continuum of decreasing phonetic explicitness" (1996: 15). The degree of explicitness is influenced by adjacent segments or by prosodic and other factors like speech rate or by language-specific or variety-specific conventions or, most likely, by a combination of some or all of these factors. Thus some processes are more "phonetically natural" than others in that they arise more directly from the physical constraints inherent in the vocal mechanism, while others must be seen as arising from cognitive processes (Nolan 1996: 19). Between the two extremes "phonetic naturalness" is a matter of degree, rather than there being a simple dichotomy between effects

7. The process-based characterisation of these phenomena implies an analysis in terms of rules operating on segments in citation forms; the discussion here will adopt that descriptive convenience, following Nolan and others, but this should not be taken as representative of a commitment to any theoretical analysis in such terms.

resulting, “from the mind or from the mouth” (Nolan 1996: 17). Of course, phonetically natural processes may also be overridden even in very rapid speech, a choice which must be cognitive, so there are evidently interactions between levels of constraints.⁸ There is no reason why there should not be abstraction from phonetic continua to discrete phonological categories, provided a case can be made for such analysis, but in the absence of a watertight case for (t,d) (see above) the aim here is to determine conversely whether there are parallels between the behaviour of word-final (t,d) stops and that of other word-final stop consonants, as characterised in terms of points along the CSP continuum, or whether (t,d) does in fact merit the special status accorded to it in variationist sociolinguistic analyses.

A comparative analysis is not an entirely straightforward undertaking, since there are some structural obstacles to direct comparisons between word-final consonants. Comparison with non-cluster /t/ and /d/ has to take account of the fact that acoustic cues are available for postvocalic consonants which are not present for /t/ and /d/ in clusters, such as formant transitions into closure from the preceding vowel. Clusters involving other word-final stop consonants are more limited in distribution than (t,d) clusters: they are always tautomorphemic with the preceding consonant; /g/ never occurs in word-final clusters; /b/ occurs in a very few, rare lexical items preceded by /l/; /p/ and /k/ are only preceded by /l/, /s/ and homorganic nasals. However, it should be noted that monomorphemic (t,d) also occurs almost exclusively with preceding /s/, /l/ and homorganic nasals (94% of the tokens analysed by Tagliamonte & Temple (2005) and 95% of tokens in the ‘demographic’ part of the British National Corpus;⁹ see Temple, ms: Tables 2 and 3), other consonants appearing mainly or exclusively in past-tense verb forms (accounting for about 28% of the total number of tokens in Tagliamonte & Temple and less than 15% of all the BNC (t,d) tokens).¹⁰ Nevertheless, with these *caveats* in mind, some useful comparisons can be made.

For convenience, the discussion will be structured round an adapted version of Nolan’s (1996) classification of CSPs, expanding it to include other combinatorial properties of word-final consonants which might be considered as leaving the

8. Nolan points out that both variable phonetic explicitness and phonetic naturalness are continua. In order to avoid confusion in the following discussion, I shall use the terms ‘scale’ to refer to the continuum between physiologically constrained and cognitively governed CSPs and ‘continuum’ to refer to degrees of phonetic explicitness. Neither continuum is truly unidimensional, as Nolan acknowledges.

9. The BNC spoken corpus is described in Crowdy (1995); the figures here are taken from the word-frequency list provided by Kilgariff and downloaded from <http://www.kilgariff.co.uk/bnc-readme.html> on January 7th, 2011.

10. Total Ns = 1118 and 78726 respectively.

essence of the segment in tact, such as [t^h] vs. [t] vs. [t̥]. I thus examine in turn release characteristics, lenition, glottalisation, voicing assimilation, place assimilation and coalescence, although the boundaries of classification are far from clear-cut, and this will be evident throughout. The analysis is qualitative: once one focuses on phonetic detail in specific contexts, numbers of tokens per cell fall to a level where it is not possible to use the kinds of statistics which can be performed on a categorical binary alternation ([t,d] vs. zero) across aggregated contexts (such as ‘before obstruents/nasals’). It is not the proportion of tokens concerned which is central to the present argument, but whether the range of realisations present in the data corresponds to that predicted by a CSP analysis of (t,d).

2.1 Release characteristics

Prepausally and prevocally, alveolar stop reflexes of the York (t,d) consonants show the range of release characteristics one might expect to find in British English: unreleased (prepausally) and released more or less strongly, /t/ with and without aspiration, /d/ sometimes devoiced. We shall not dwell further on prepausal or prevocalic tokens in this sub-section. It is no surprise that rates of deletion of non-prepausal (t,d) consonants across studies have consistently been found to be considerably higher before consonants than before vowels,¹¹ and highest before other stops, where they are least likely to be released audibly. This effect would rank very much towards the phonetically natural end of Nolan’s “mouth-mind” scale. Nevertheless, logically if there is stop closure this has to be released somehow in order to articulate any following sound, including consonants. Henderson & Repp (1981) examined word-internal heterosyllabic and word-final tautosyllabic stop sequences in read speech. On the basis of acoustic analysis and perceptual tests they propose a five-point scale of phonetic classification of stops: unreleased, silent-released (no clear acoustic burst), inaudible release (clear acoustic evidence of a weak burst, but imperceptible), weak release, strong release. They did not test C.C sequences across word boundaries, but suggest that the word-internal condition (where the consonants were generally heteromorphemic as well as heterosyllabic) is somewhat comparable, so one might expect to find the same range of effects. The articulatory and aerodynamic conditions affecting the second consonant in a word-final cluster are, of course, different but it remains the case that where there is consonantal closure there will have to be separation of the

11. In African American Vernacular English (e.g. Wolfram 1969) the difference can be much less, but these varieties also show patterns of social stratification (particularly pre-vocally) which are generally not found elsewhere and arguably cluster reduction here is a truly sociolinguistic variable and not just the effect of a combination of CSPs.

articulators in order for a third, word-initial consonant to be produced. Before looking at cluster-final consonants followed by stops in the more naturalistic York data, we first examine the range of releases found there in word-final singletons in the same context.

Various types of release¹² occur in word-final singleton /t/ and /d/ followed by stops, though the limited distribution of word-final /t/ and the preponderance of glottalised realisations, particularly in the highly frequent words where it most often occurs (e.g. *it*, *got*), makes examples of voiceless final alveolars harder to find.¹³ There are nevertheless examples of clearly released [t], as in (1):

- (1) *and hot coals* [hʊtkʰʊʊlz] *used to drop out*

and of clearly articulated [tʰ] with no acoustic or auditory evidence of release, as in (2):

- (2) *another catch would detect that you'd got eight bales* [ɛɪtʰbeɪlʰz]

as well as less clear examples of unreleased voiceless stops whose place of articulation is difficult to determine, as in (3), where the very short preceding vowel and glottal reinforcement make it hard to tell whether the word *cut* ends in a [tʰ] or a [pʰ] assimilated to the following [m]:

- (3) *they cut my* [tʰʊɪ̯tʰmə] / [tʰʊɪ̯pʰmə] *trousers off me*

/d/ occurs in a wider range of lexemes and shows all types.¹⁴ Examples (4) and (5), where the following consonant is /m/, illustrate the same sequence of words uttered by the same speaker in the same stretch of discourse (talking about traditional Morris dancing), with the word-final /d/ weakly released in (4) and unreleased, with no acoustic burst, in (5):

- (4) *So we do Escrick which is long sword metal* [sɔ:dmetl̩]

- (5) *there's long sword metal* [sɔ:d̥metl̩:]

There are also a few examples by this and other speakers of inaudible release accompanied by a clear, if weak, acoustic burst, as in (6):

12. We make no distinction here between Henderson & Repp's first two categories (unreleased vs. silent release with no acoustic burst), since none of the tokens discussed are in absolute final position. Neither is it necessary here to make a systematic distinction between weak and strong audible release, although the presence or absence of audible aspiration is noted in the transcriptions.

13. For example, the recording of SW, who produced (1) and (2), contains 48 tokens of word-final singleton /t/ followed by stops, 39 of which were in frequent function words.

14. Speaker SW produced fifteen tokens of singleton /d/ before stops, including (4) and (5).

(6) *and I never did get [dɪdʒəʔ] round to seeing it*

Word-final stop consonants at other places of articulation are rarer¹⁵, but cases of both released and unreleased articulations are found with following stops, as in (7)–(8) and (9)–(10) respectively, and there are even examples of inaudible release with a weak acoustic spike, as in (11), which is illustrated in Figure 1:¹⁶

(7) *my grandfather used to go to a pub down [pʊbdaʊn] there*

(8) *there's a lot of (...) sick people [sɪkpiːpl] as in...*

(9) *followed the the cop car [kʰɒpˈkɑːɹ]*

(10) *and you roll it up into a big ball [bɪɡˈbɔːl] and stick it on the end*

(11) *primary school goes from reception up to [ʊptʰə] year 6*

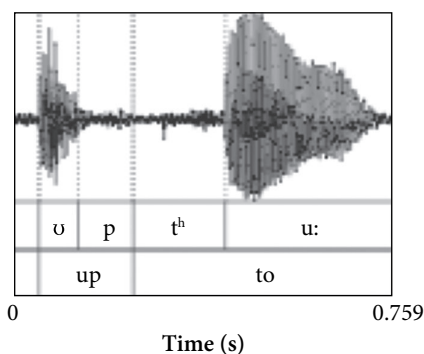


Figure 1. Waveform showing *up to* [year six] (11) with inaudibly released [p]; female speaker.

As for the (t,d) cluster tokens,¹⁷ there are no cases of inaudible release with clearly visible bursts, but both released and unreleased reflexes of both consonants may be found before following stops. Many of the released tokens occur when the speaker is hesitating, as in (12), or pausing for a discourse effect, as in (13):

15. SW has eighteen pre-stop singleton tokens of /p/, fourteen of them in the word *up*, sixteen tokens of /k/, six of /g/ and none of /b/.

16. Space precludes illustration of every example so a small selection is provided here. Spectrograms and sound files of all examples are available on the website accompanying this volume at <http://dx.doi.org/10.1075/silv.15.media>

17. (t,d) tokens are taken from the original analysis in Tagliamonte & Temple (op.cit.), which was selective in order to maximise even distribution across speakers, morphological classes and lexical items. The average number of tokens per speaker with a following stop was 8.6.

- (12) *he'd left [lɛft̚] (.) Betty with nothing*
 (13) *and he found Minesweeper [faʊnd̚] (.) [ma:ɪnswi:p̚ə], have you played Minesweeper?*

But there are also clear cases where no pause is involved, as in (14) and (15):

- (14) *like my hands would have been fucked basically [fɔrt̚bɛsɪklɪ]*
 (15) *in an underground bunker [ʊndəɡraʊndbʊnk̚hə]*

(16) and (17) show unreleased /t/ and /d/ respectively:

- (16) *your needles left particles [lɛft̚pa:tɪkl̚vz] in the groove of the record*
 (17) *been told by [tɔ:ld̚baɪ] that many people*

In non-(t,d) clusters the same range of patterns is found, albeit to a much lesser extent, as illustrated by (18) and (19):

- (18) *I'm trying to think now [θɪŋknaʊ] how I can make...*
 (19) *just don't ask me [ask̚mi] for help*

These examples demonstrate clearly that coronal-stop reflexes of (t,d) consonants exhibit the same range of realisations as other singleton and cluster-final plosive consonants when followed by a stop in connected speech. This observation on its own poses no problem for the generally accepted account of (t,d), but we now turn to some rather more problematic issues for that account.

2.2 Lenition

In this section I first compare the range of lenition patterns in (t,d) with that in the comparator word-final consonants, then examine the possibility that there are sociolinguistic constraints on (t,d) which might differentiate it from other cases of full lenition at word boundaries; I next assess whether the contextual influences on full (t,d) lenition are consistent with a CSP analysis or require specific phonological rules, and finally I identify cooccurrence patterns with lenition of other consonants in a given string.

2.2.1 Lenition patterns in word-final stops

In his commentary on Nolan's (1992) discussion of alveolar-to-velar place assimilation, Hayes proposes a general phonetic rule of word-final alveolar weakening, on the grounds that, "[f]or example, the segment /t/ is often weakened in its articulation even when no other segment follows" (Hayes 1992: 284). In fact, very few of the unambiguously realised (t,d) consonants are weakened alveolars in the York data, but there is some evidence of the expected "continuum of phonetic

explicitness” whether or not another consonant follows. Examples (20)¹⁸ and (21) show somewhat lenited prevocalic /t/ and /d/, the latter also being devoiced, along with the preceding and following segments, and (22) shows a rather greater degree of gestural weakening, to a retracted fricative articulation. (20) is illustrated in Figure 2.

(20) *(it) was the discipline I liked and [lɛɪt̚ˈsɛn] that was all there was to it*

(21) *she wa’n’t very pleased wa’n’t [pɫ̚ˈɪz̥t̚ˈhʌvən?] my mum*

(22) *they went and knocked on [nɒˌʒɔn] Andrew’s door*

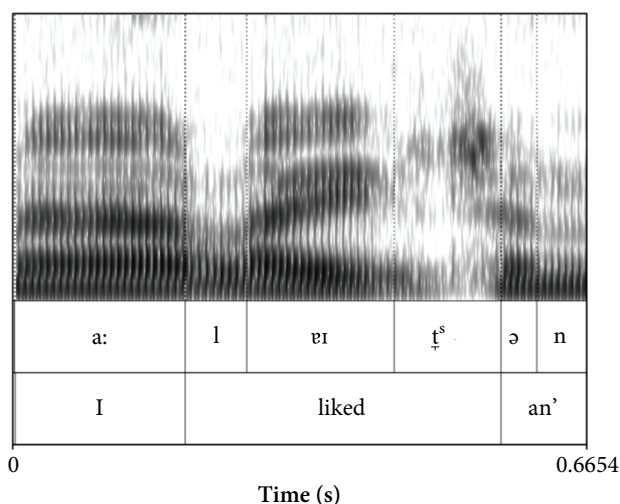


Figure 2. Spectrogram showing *I liked an'* (20); male speaker.

Parallel examples of lenition are found with singleton /t,d/ and other stops but again these are relatively uncommon. (23), illustrated in Figure 3, is a very lax, slightly fricativised articulation:

(23) *it really reminded me [ɹ̥ˈmʌndɪd̥ˈwɛɪ]*

18. There is some debate in the literature (Buizza 2010, 2011) as to whether affricated release constitutes lenition or fortition, but the York data seem similar in this respect to the alveolar affrication found in “Modern RP” by Buizza and to be further instances of lenition, often co-occurring with a lenited stop articulation, as here.

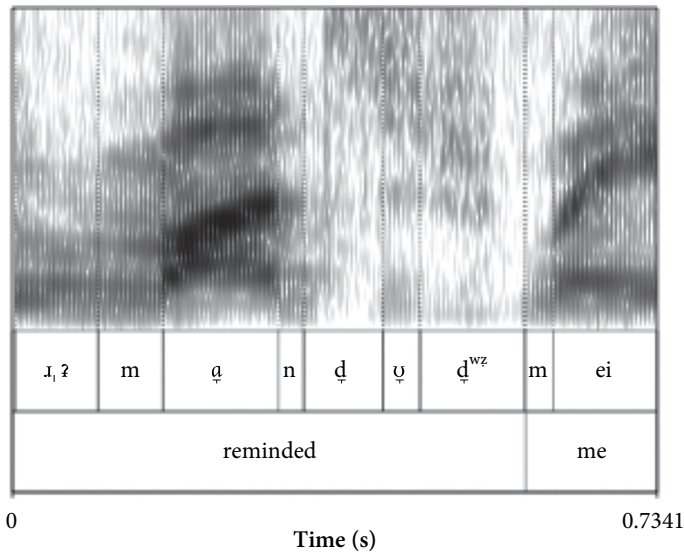


Figure 3. Spectrogram showing *reminded me* (23); female speaker.

Examples (20) to (22) would count along with non-lenited stops as non-applications of a variable rule of coronal stop deletion, and that is how they were treated by Tagliamonte & Temple (2005). The deletion rule would be said to have applied only at the extreme open end of the continuum of lenition, where there is no residual auditory or acoustic evidence of a reflex of /t/ or /d/. Once again we find parallel cases: there are examples of fully lenited word-final singleton consonants, as in (24), which is very rapid speech, and (25), where the vowel preceding the deleted consonant is stressed and lengthened, indicating that this full lenition is not necessarily dependent on a rapid speech rate:

(24) *they had the coal delivered by* [dɪlɪvəbə] *rail*

(25) *and it was very vague because* [vɛɪbɪk^hʊz]

and examples of full lenition in non-(t,d) clusters:

(26) *and my grandchildren are able to help* [təʃel^v]¹⁹

(27) *they didn't ask me* [ˈʔasˈmi] *so...*

19. This token has no trace of labialisation, despite the fact that it is followed rapidly by an inbreath and the word *When*, beginning a new sentence.

Word-internally, deletion is probably lexicalised in most cases, occurring nearly categorically in words like *grandmother*, *grandfather* and *Christmas*, but it also occurs in less frequent compounds, such as *landmarks* (28) and *second-hand* (29):

(28) *one of the local landmarks* [lɔ:kɫʰ:anma:ks] *was this brickyard chimney*

(29) *they bought things in second-hand shops* [sɛŋ'hantʃɒps]

As with (t,d), most, though not all, of these examples are pre-consonantal and so an unsurprising outcome of “phonetically natural” CSPs. Indeed Nolan (1992) gives a hypothetical example of the total lenition of word-internal /d/ in the word *hundred* (“[hʌndɹɪəd] (? → [hʌndɹɪəd]) → [hʌnɹɪəd]”; Nolan 1992: 23), which he classifies under “Target Undershoot” at the phonetically natural end of his scale. Is there, then, any independent evidence that word-final (t,d) clusters are quantitatively or qualitatively different from Examples (24) to (29), which would justify their treatment as a special variable rule?

2.2.2 Sociolinguistic variation in lenition

One type of evidence for the special treatment of (t,d) would be sociolinguistic effects not applying to other cases of word-final lenition/deletion. Such effects have been found for AAVE and some southern US dialects, but not for other varieties of English. Gimson’s classic text on English pronunciation (as re-edited by Cruttenden) is peppered with what are essentially sociolinguistic judgements, such as the comment that, “the elision of one of a boundary cluster of only two consonants sometimes occurs in casual speech but is usually characterised as sub-standard, e.g. *He went away* /hɪ wɛn əˈweɪ/ (...) *Let me come in* /lɛmɪ kʌmˈɪn/” (Cruttenden 2008: 302). Interestingly, where word-final clusters are concerned, these contexts correspond exactly to the prevocalic cluster reduction noted as a qualitative and quantitative sociolinguistic difference between African American and other varieties of North American English. By contrast, it is striking that (a) word-final clusters are grouped in the above quotation with cross-boundary sequences, suggesting no special status, and (b) no such evaluative judgements are proffered in comments on the deletion of cluster alveolars before consonants, which Gimson/Cruttenden seem to treat as straightforward, socially unmarked CSPs, entirely to be expected in RP:

[...] sounds may be elided in fast colloquial speech, especially at or in the vicinity of word boundaries (...) In addition to the loss of /h/ in pronominal weak forms and consonantal elisions typical of weak forms, the alveolar plosives are apt to be elided. Such elision appears to take place most readily when /t/ or /d/ is the middle one of three consonants. (Cruttenden 2008: 303)

[...] Where the juxtaposition of words brings together a cluster of consonants (particularly of stops), elision of a plosive medial in three or more is to be expected, since because of the normal lack of release of a stop in such a situation, the only cue to its presence is likely to be the total duration of the closure.

(*ibid.*: 304)

Consistent with this observation, (t,d) in the quantitative analysis of York data was not found to pattern with independent social variables, except for a weak tendency for male speakers to delete more frequently than females (Tagliamonte & Temple 2005: 296–297).

2.2.3 Contextual effects on full lenition

So far as linguistic constraints on the variability were concerned, Tagliamonte & Temple found a very strong effect of following phonological segment, with deletion highly favoured before following consonants and disfavoured before following vowels, as had all previous studies. Gimson/Cruttenden's account of elision/deletion, just cited, and similar patterns found in other languages such as Dutch (e.g. Schuppler et al. 2009) suggest that this is more likely to be the result of variable CSPs than of a specific variable phonological rule. The more detailed distributional effects are consistent with this interpretation: following obstruents and nasals favour deletion more than glides and liquids. A further breakdown of the data is presented in Table 1, which shows the results of a multivariate analysis of the effects of preceding and following phonological segment using GoldVarb (Sankoff et al. 2011). The factor weights assigned to following nasals, stops and fricatives appear to justify their treatment as a single statistical factor, which is the common practice with this variable; however, /h/ is here separated from the other following fricatives and clearly behaves very differently. In fact, over half the tokens with following /h/ actually have a following phonetic vowel and the rates of deletion are identical in these tokens and those with following [h] (10% vs. 11%). This again is consistent with a CSP analysis of (t,d), showing that it is following consonants with close oral constriction which inhibit overt reflexes of /t,d/, whereas /h/, with glottal constriction but more open oral articulation, patterns more like vowels.

Although the quantitative results appear to confirm that the CSP analysis of (t,d) is reasonable, the causal link is not so straightforward since, as indicated in the comments on following /h/, they follow the convention of analysing the phonological context in terms of the underlying representation. This practice is consistent with the view of (t,d) as a variable rule which applies in the lexical phonology as well as post-lexically, but it poses analytical problems such as the relative ordering of this and other rules affecting particularly the preceding phonological context, for example *l*-vocalisation. These problems are discussed in more detail

Table 1. Results of GoldVarb analysis of the effects of following and preceding phonological context on deletion of /t,d/.

	Factor weight	% deletion	Total N
FOLLOWING CONSONANT			
nasal	.918	70	69
stop	.890	66	93
fricative	.887	62	101
glide	.690	38	106
/r/	.605	28	29
/l/	.496	25	24
/h/	.354	11	62
vowel	.291	8.3	507
pause	.200	5.5	127
RANGE	[72]		
PRECEDING CONSONANT			
/s/	.690	41	303
/ʃ/	.565	31	64
nasal	.497	21	329
stop	.382	16	169
liquid	.374	21	126
non-sibilant fricative	.298	12	127
	[39]		
TOTAL			1118

in Temple (ms). They are far less problematic when (t,d) is analysed in terms of CSPs, so we once again turn to the qualitative data to confirm whether this non-phonological analysis can be justified.

I present here just a small sample of typical (t,d) tokens with different combinations of preceding and following consonants, where the variable rule analysis would state that deletion has applied, beginning with cases where the preceding and following consonants are pronounced in their unlenited citation forms. Examples (30)–(33) are typical of target undershoot in continuous speech:

- (30) *oh I'd booked my [bukˈmʰ] ticket, yes*
- (31) *but we still kept corresponding [kʰɛpˈkʰɔɪspɔɪndɪŋ] all the time*
- (32) *so of course I left school [lɛfskuˈlʰ] at fourteen*
- (33) *... whether I spent the first few [fɜːsfjuˈw] months of my life*

It is not necessary to assume here that the speaker has deleted the (t,d) consonant in the phonology and therefore produces no alveolar closing gesture; rather, it is perfectly plausible that these are cases where the hypothetical target for the /t/ or

/d/ is, “attained less completely in phonetically less explicit pronunciations” (Nolan 1992:23). Such undershoot is not solely a function of the segmental context, as shown by the lenition of word-final singletons in (24) and (25), but as pointed out in the quotation from Cruttenden (2008) above, it is especially to be expected in sequences of three consonants, particularly with stops. This is not always the case (see (14)), but nineteen of the twenty-two tokens in the York data set with both preceding plosives and following plosives or nasals are elided. Similarly, the progression from fricative to fricative without an intervening stop articulation in (32) and (33) is normal in fluent (British) English. In fact, only four out of a total of 71 York (t,d) tokens with both preceding and following fricatives have any audible or acoustic phonetic reflex, and those are all preceded by the voiced weak fricative /v/. These effects are compounded when preceding and following consonants share both place and manner of articulation, as shown in (34), where there is a fluent transition from [p] to [m] with the speaker maintaining the bilabial closure:

(34) *they stopped making* [stɒpˈmeːkɪm] *bricks er yonks ago*

In (35), where the manner of articulation is different but place is labial in both consonants, the elision is again unsurprising, with a fluent transition from labio-dental constriction to bilabial closure:

(35) *'think that's what saved my* [se:vɪmə] *back*

In some cases, the preceding consonant is slightly lengthened, which might be construed as cueing the underlying coronal segment, as in (36) and (37):

(36) *and we were kept busy* [kɛpˈˌbɪzi:]

(37) *only when I left school* [ˈlɛfˈskuːl]

However, this is not always the case, and indeed evidence for a direct link between closure duration and the number of underlying consonants is equivocal, as confirmed by Kühnert & Hoole (2004), whose articulatory data obtained from electromagnetic articulography (EMA) showed that “the complete fusion of two velar stops in fast speech could (...) result in closure durations identical to an individual stop (...), a healthy reminder that the interpretation of closure duration in fluent speech still has to proceed cautiously” (Kühnert & Hoole 2004: 572).

In all the cases of deletion, there may, as indicated by Gimson/Cruttenden, be a residual alveolar gesture indicating that from a production point of view the (t,d) consonant is somehow present. This could involve a lenited gesture resulting in the uninterrupted frication of (37) or full contact masked by the maintenance of bilabial constriction in, e.g., (34) and (36). Note, however, that the (t,d) cases are not unique in this respect: it is perfectly possible that gestural overlap might

have occurred in *delivered by* in (24) resulting in the percept of deletion despite a full or lenited alveolar gesture for the final singleton /d/. We return to the matter of such residual gestures in the discussion of assimilation in §2.5.

Relative timing of gestures may also account for deleted tokens with preceding sonorants. In (38) there is coronal closure for the preceding lateral consonant; it is possible that the sides of the tongue were raised before the release of this closure, essentially forming a [d] or [ɖ]:

(38) *but there was all old carpets* [ɔlk^hap^hiɾts] *and pictures*

In (39) the timing overlap is between the transition from alveolar to bilabial closure on the one hand and the raising of the velum for the cessation of nasality on the other:

(39) *something like eight thousand people* [θauznp^hi:pəl^y]

Example (39) contrasts with (40), where nasality ceases before the bilabial closure:

(40) *they were rather like unmanned bombs* [ʊnmænd^hbɒmz]

Examples (39) and (40), which are illustrated in Figures 4a and 4b, are directly analogous to Nolan's hypothetical continuum for *hundred* (see above, p. 106), suggesting again that the most straightforward account of deletion or not here is a CSP one.

In similar vein, CSPs towards the natural end of the scale provide a straightforward account of deletion between nasals. The velum is known to move more slowly than other articulators. It would therefore require extra articulatory effort to produce (41) with an oral [d] closure (released or unreleased) between the alveolar and nasal preceding and following consonants:

(41) *then it'll have locked behind me* [bɪfɪɛnmɪ]

2.2.4 Co-occurring patterns of lenition

Finally, if these cases of lenition are a function of general CSPs on a continuum of decreased phonetic explicitness, one would expect that they would co-occur with characteristics of lenition in other segments, and this is indeed what we do find. In (20) and (22), above, lenition of final /t/ is accompanied by lenition of the first consonant of the cluster, /k/, which in (20) is fully elided and in (22) is realised with glottalisation, in the form of creaky voice, but no acoustic evidence of a velar gesture. *l*-vocalisation is not a common feature of the York dialect, but there are ten tokens (out of 130) where /l/ in (t,d) clusters is vocalised, and some where it is elided altogether. When there is a following word-initial consonant, these always

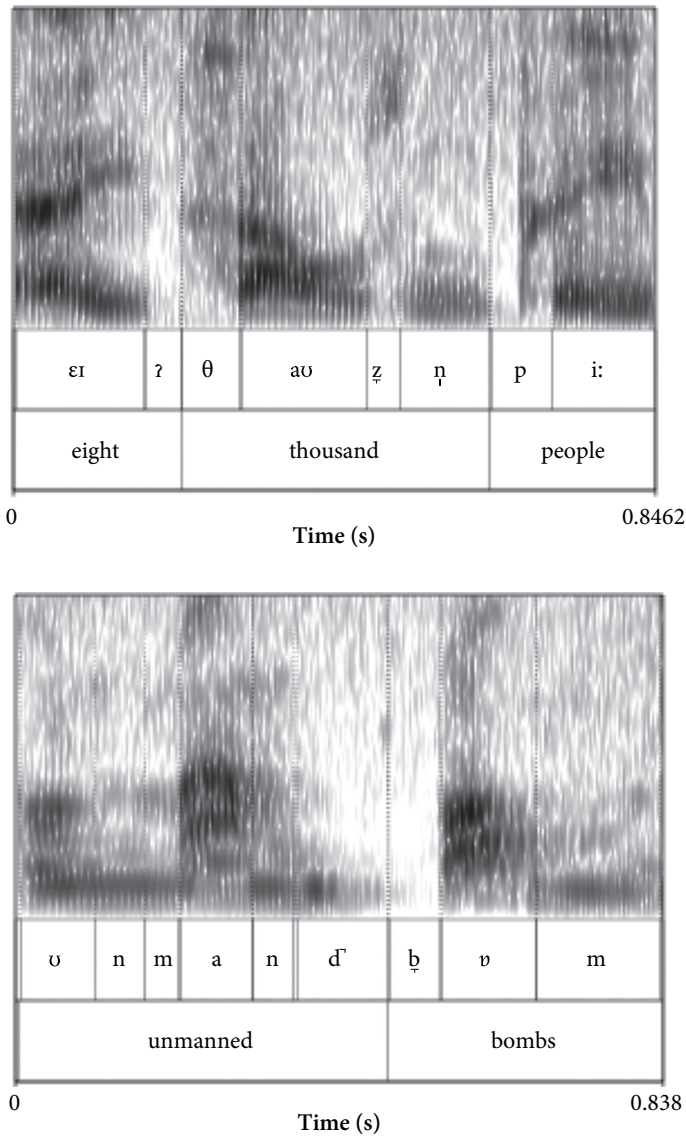


Figure 4. Spectrograms showing (a) *eight thousand people* (39); male speaker and (b) *unmanned bomb(s)* (40); female speaker.

co-occur with /t,d/ deletion, as in (42). *told* here is unstressed and spoken very fast; as well as deletion of preceding /l/, the following vowel is reduced and the /t/ of the following preposition is also lenited. The stressed /ld/ cluster of *hold*, by contrast, has both preceding [l] and released [d]. Here the cumulative evidence suggests that the deletion of /d/ is simply one of a set of cooccurring CSPs, which are a function of speech rate and accentual patterns.²⁰

(42) *and they said (.) told me to [tʰʊmɪdə] hold it*

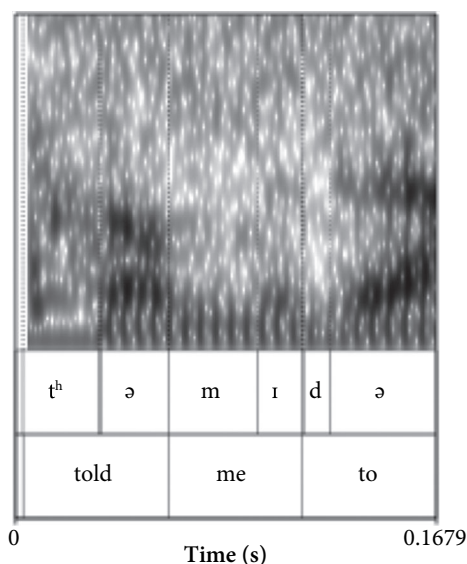


Figure 5. Spectrogram showing *told me to* (42); male speaker.

Viewed through the lens of lenition, then, the behaviour of (t,d) preconsonantly shows a range of decreasing phonetic explicitness paralleled in other word-final singleton and cluster consonants, and explicable in terms of lenition partly as a function of the surrounding phonetic context. In British English, at least, (t,d) lenition shows no distinctive sociolinguistic patterning and it is seen to co-occur with varying levels of phonetic explicitness in surrounding segments. We now turn to examining the interactions, as opposed to simple co-occurrence patterns, of (t,d) lenition with other well known CSPs.

20. Cf. Nolan again: "Segmental CSPs are not independent of prosodic CSPs – they are sensitive to the prosodic restructuring which the latter bring about, and ultimately may turn out to be treated best in conjunction with the prosodic changes" (Nolan 1992: 18).

2.3 Glottalisation

Hayes (1992) further comments on weakly articulated /t/ that, “in such cases, the weakened /t/ is usually ‘covered’ with a simultaneous glottal closure” (Hayes 1992:284–5). There are many instances in the York data of (t,d) tokens clearly surfacing as glottals (N=47), although these were generally not accompanied by auditory or acoustic evidence of an alveolar articulation except in a relatively small number of cases, such as the second element in the compound of (43) (the same utterance as (29)):

- (43) *they bought things in second-hand shops* [sɛʔnˠhantʔʃɒps]

The percept was most frequently as a glottal stop, but the acoustic evidence showed that the data included both full glottal stops, as in (44), and continuous glottalised realisations perceived as glottal stops or creaky voice, as described by Docherty and Foulkes (2005) and exemplified in (45).

- (44) *all the way, went all the way back because* [wɛnʔɔɪʔə]

- (45) *it really spoilt my memories of school* [spɔɪʔɪˠmaɪ]

Most were reflexes of underlying /t/, but there were a few cases of devoiced glottalised /d/ followed by a voiceless consonant, as in (43) above.

Non-cluster /t/ is very frequently realised as a glottal (46), as is /k/ (47):

- (46) *he got knocked over* [gɒʔnɒʔtˠɔʊə]

- (47) *I used to quite like bikes* [lɛəˠbaɪks]

Word-final /p/ (48, 49) and /k/ in clusters are also glottalised, the latter most frequently in *think* (50),²¹ but also in other words (51):

- (48) *she had to come and help me* [ɛɪˠwɪ]

- (49) *'cos it's finished being a training camp now* [kamʔnaʊ]

- (50) *I think we went to Scarborough* [θɪŋˠɔwɛnʔ]

- (51) *you're not supposed to take it with milk* [mɪlˠ]

Apart from one token with preceding /s/ and one with preceding /p/, both in very frequent words (*just* and *kept* respectively), all glottal (t,d) tokens were preceded by /l/ or /n/. However, this is a slightly misleading observation, since two thirds of velar stops forming the preceding phonological context (75/109) were also

21. Speaker NB, for example, produces 105 tokens of *think*. All 46 /k/s with following stops are glottals; all 47 tokens followed by vowels or /h/ are realised as [kʰ]; with a following pause there are three glottals and nine plosives.

realised as glottals, as illustrated in (46) above. This is unproblematic in cases such as (46) or (52) or (53),²² where there is a clear sequence of a glottal plus following released [t^h]:

(52) *if if a project or [pɪəʊdʒɛɪt^hɔ:] contract comes up*

(53) *and they evacuated the whole place except us [ɪ_sɛ_t^hʊʃ]*

However, in cases such as (54), where there are not two clearly distinguishable articulations, it is often impossible to determine of which underlying segment the glottal is a reflex:

(54) *if if a project or contract comes [kɒntɹ̥aʔkʊmz] up²³*

In (54) the glottal is slightly lengthened, which may possibly be taken as evidence that it is a reflex of the two consonants, but there are many other examples where the glottal is not notably long, such as (55), and as mentioned above, length is not an unequivocal indicator of the presence of more than one consonant.

(55) *She knocked straight [nɒʔstɹ̥ɛɪ_] into us yeah*

Cases such as (54) and (55) pose problems for a rule of consonant deletion conditioned primarily by the following and preceding phonological context: should the glottal in any given case be taken as the reflex of the preceding consonant or the (t,d) coronal stop or both? how might one decide the correct analysis in each case? The answer to these questions determines whether or not the (t,d) consonant is deemed to have been deleted. These and related questions are discussed by Temple (ms) as methodological/analytical problems for the treatment of (t,d) clusters with a categorical deletion rule. In the light of the present discussion, viewing the behaviour of the clusters as the expected result of variable CSPs would seem to provide a coherent alternative analysis. Variable glottalisation of any voiceless stop is context-specific and dialect-specific, and known to be a sociolinguistically changing feature of British English (e.g. Fabricius 2002; Foulkes & Docherty 2005; Stuart-Smith et al. 2007) and therefore must feature in the speaker's cognitive phonetic plan. Moreover, as Nolan points out, glottalisation cannot be seen as phonetically natural lenition, since it involves increased constriction of the glottis, "an articulation in direct conflict with the opening gesture required for [t^h] (or any other non-glottalised stop)" (Nolan 1996:21). As we have seen, it is normal in this variety for all final stops and for

22. (53) is the only case of glottalised preceding /p/ in the York (t,d) data set.

23. (52) and (54) represent the same utterance and reproduce (3) and (7) from Temple (ms).

penultimate /k/²⁴ be realised as glottals; the pertinent variability would seem, then, to be between glottalised and non-glottalised realisations of final clusters, rather than between C₂ alternating between zero on the one hand and [t] or [ɾ] on the other, with word-final (cluster) codas that consist only of a glottal stop somewhat arbitrarily being deemed as having a deleted or undeleted /t/. Whether or not these glottal-only codas simultaneously ‘cover’ a weakened (or indeed non-weakened) alveolar articulation is unknowable from auditory/acoustic data alone, but all except one of the 55 (t,d) tokens with ‘preceding’ glottals and following vowels or pauses have alveolar release, which suggests that some alveolar articulation could be present preconsonantly too. Any alternation between the presence or absence of a ‘covered’ alveolar gesture in glottal-only codas may well be a combination of idiosyncratic (and therefore cognitive) and physiological constraints (target undershoot). And the presence of an observable release before almost all vowels and no stop consonants, and before four out of nine following continuants is towards the natural end of Nolan’s scale. The behaviour of all glottal codas would appear, then, to be a function of a combination of both cognitive and more ‘natural’ CSPs. Once again this observation is reinforced by the co-occurrence of glottalisation with other CSPs, as in (56), with its fully lenited nasal, which is illustrated in Figure 6:

(56) *they went and [ðɪwɛ̃n] knocked on Andrew’s door*

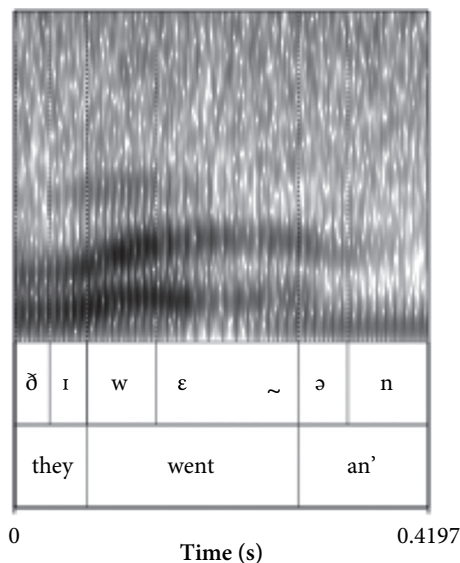


Figure 6. Spectrogram showing *they went an'* (56); male speaker.

24. This occasionally also applies to /k/ before plural /s/ as in *I've only done it for three weeks* [wiːz̥s].

2.4 Voicing assimilation

There is no assimilatory voicing of voiceless (t,d) to following voiced consonants in the York data set, although there is at least one token of partially voiced /k/ in a cluster, which is shown in (18) above. By contrast, most released tokens of /d/ are devoiced by assimilation to a following voiceless consonant, as in (57) and (58):

(57) *how I can make an old-fashioned copper* [faɪnḁkʰɔpʰə]

(58) *there was a lot of old people* [əlʰtpʰi:p]

This is as might be expected from the well known phenomenon of Yorkshire assimilatory devoicing, although York English seems to show gradient devoicing rather than the categorical neutralising devoicing described by Wells, where “*wide trousers*, having undergone Yorkshire Assimilation, is a perfect homophone of *white trousers* [ˈwaɪt ˈtraʊzəz]” (Wells 1982: 367) and it is clearly different from the categorical assimilatory glottalisation in the West Yorkshire variety studied by Broadbent, where “the /d/ never surfaces as a [t], as one might expect, so ‘vodka’ *[vɒtkə] and ‘godfather’ *[gɒtfa:ðə] are impossible realisations” (Broadbent 1999: 19).²⁵ This gradience is also evident in the singleton consonants in (59) and (60), illustrated in Figure 7:

(59) *it was a lad called* [lɑdʰɒlʰd] *Wayne*

(60) *choose to be a good friend* [gʊtʰfɛɔ.ŋ:t]

In clusters, the devoicing can extend to the first consonant, as in (61), and this can apply in cases of apparent deletion like (62):

(61) *so she’s moved quite a* [muːytkwɛɪɔ̥] *way away*

(62) *he actually lived seven* [lɪy:sɛvən]

The juxtaposition of these two examples shows once again that coarticulatory phenomena affecting the first consonant of the cluster cannot necessarily be taken to indicate the deletion of the second. More importantly, here again we have unsurprising CSP patterns both with (t,d) clusters and with other singleton and cluster word-final stops.

25. In fact, the only (t,d) token with a lexical /d/ realised as an assimilatory glottal is the glottally reinforced final consonant of *second-hand* shown in (43) above.

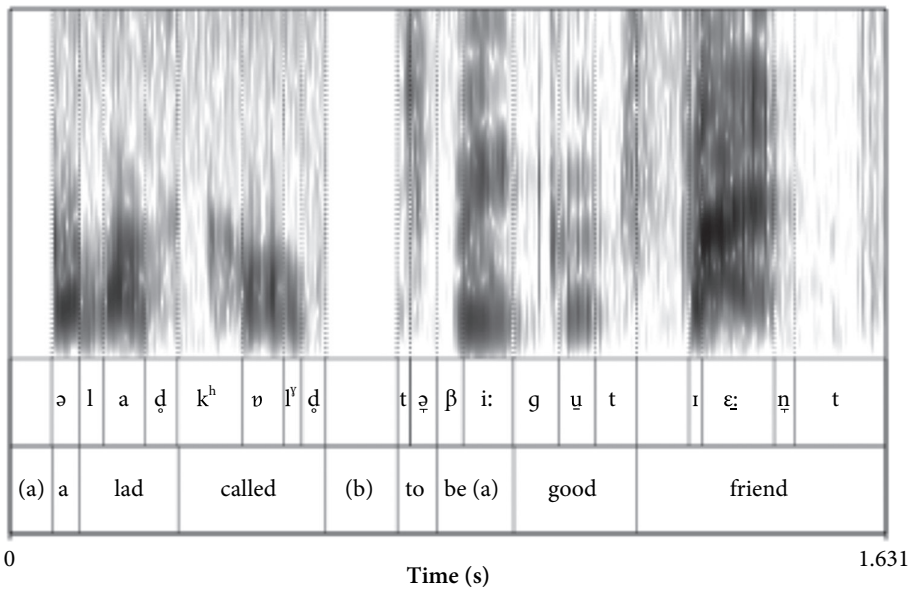


Figure 7. Spectrograms showing (a) *called* Wayne (59) (b) *to be a good friend* (60); female speakers.

2.5 Place assimilation

Assimilation of place has become a central topic in discussions of the relationship between phonetics and phonology in the wake of numerous studies examining gradience versus categoricity, particularly with reference to residual gestures (e.g. Barry 1985; Nolan 1992; Ellis & Hardcastle 2002; Kühnert & Hoole 2004; Bermúdez-Otero 2010b). It is well known that in English, “word-final /t d n s z/ readily assimilate to the place of the following word-initial consonant” (Cruttenden 2008: 301) but there are no clear manifestations of this in the York (t,d) tokens. The very few examples which might be interpreted this way are of glottalised tokens with preceding /n/ produced with a lengthened bilabial nasal assimilating to a following bilabial, as exemplified in (63) and Figure 8.²⁶ As already noted, however, length is an unreliable indicator of multiple underlying segments, although the qualitative change in the creaky voice suggests there may be oral reflexes of both /n/ and /t/ here.

²⁶. This token was excluded from the statistical analyses reported in Tagliamonte & Temple (2005) for reasons explained in that paper.

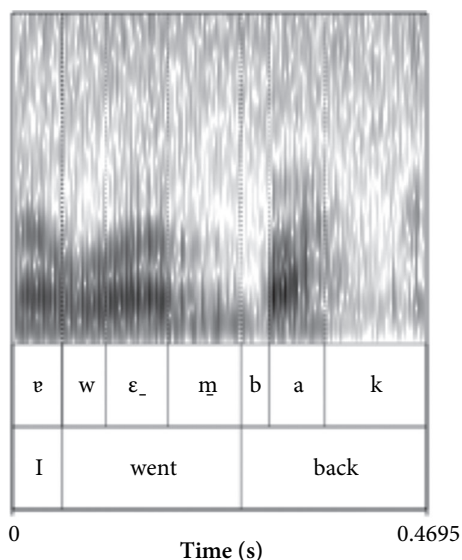


Figure 8. Spectrogram showing extract from *and then I went back* (63); female speaker.

(63) *and then I went [wε̃m] back to work again*

Singleton alveolars assimilate fairly frequently to following bilabials and velars, as in (64)–(66), although this seems to be limited to certain individual speakers and there are plentiful examples in preceding sections of non-assimilated tokens. This shows, nevertheless, that regressive assimilation of alveolars is a feature of this variety of English.

(64) *the next morning they were all brought back [bɔ:ɹp̃ba^hk̃] again*

(65) *they were really like sad people [sab̃pipl̃] straight up*

(66) *and my leg could move [k̃^hʊb̃mu:v]*

The absence of assimilation in (t,d) tokens is in fact not so surprising when the phonetic details of the data are considered. There are clearly non-assimilated alveolar stop articulations illustrated in §2.1, but there are many tokens where it was impossible to determine the place of articulation of the (t,d) consonant because of the absence of formant transitions into and out of the closure, due to the presence of the preceding and following consonants. An example with preceding /l/ is given in (67) and Figure 9 which is acoustically and auditorily ambiguous.

(67) *we've been told by [t̃^hɔ:ld̃ba:] / [t̃^hɔ:lb̃ba:] that many people*

Glottally reinforced tokens are equally difficult to identify even in singleton consonants, as illustrated by (3) above, which is reproduced here for convenience:

(3) *they cut my* [^hʊɪ̯t̪mə] / [^hʊɪ̯p̪mə] *trousers off me*

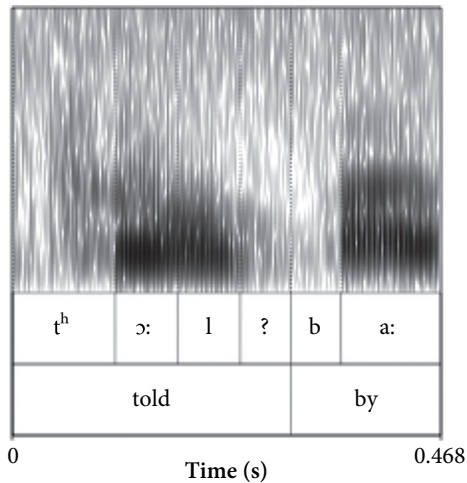


Figure 9. Spectrogram of *told by* (67); female speaker.

Moreover, fully glottal realisations of both (t,d) consonants and their preceding stops might not only be masking a possible residual alveolar gesture, as noted in §2.3, but also any assimilatory gesture which might be present, as in (68), from the same sentence as (41).

(68) *then it'll have locked behind* [lɔɪ̯ˈbɪfɦeɪn] *me*

The presence of assimilation in the York (t,d) data is much easier to determine when it involves the preceding consonant, as predicted by Gimson/Cruttenden: “When alveolar consonants are adjacent in clusters or sequences susceptible to assimilation, all (or none) of them will undergo assimilation” (Cruttenden 2008:302). However, although this is certainly true of all the unassimilated examples presented in this paper, the assimilation of preceding consonants has the consequence of rendering the word-final consonant difficult to identify and there are no tokens in this data set with assimilated penultimate and word-final consonants both unambiguously present. Instead we find assimilating preceding consonants in cases of apparent deletion, which may well be masking residual alveolar gestures, just as Browman and Goldstein found for *nabbed most* [næbmɔ:st] and *seven plus* [sɛvɪ̯plʌs] in their study of X-ray microbeam data (Browman & Goldstein 1990:365–367). This is perhaps unsurprising, since there is evidence that alveolar nasals are more susceptible to assimilation than stops (Hardcastle 1994) and most assimilated preceding consonants in the York data are nasals, as illustrated in (69) and (70), although there are also some assimilations involving preceding /s/, as in (71):

- (69) *a a sound box* [saʊmbɒks] *was only a diaphragm*
 (70) *we built, um, Bradford combined court* [kʰəmbaɪŋkʰɔːʔ] *centre*
 (71) *went to Ireland last year* [lafjiʔ] *fishing*

There are very few articulatory studies of assimilation in word-final clusters as opposed to singleton consonants, so it is not clear whether assimilated CC# clusters exist in quite so clear-cut a way as suggested by the hypothetical examples provided by Cruttenden (e.g. “*He won’t* /wəʊŋk/ *come* [...] *He found* /faʊmb/ *both, a kind* /kaɪn/ *gift”*; Cruttenden 2008:302). It would be possible to disambiguate cases where the (t,d) and following consonant differ in voicing, from auditory/acoustic data but it is difficult to see how to decide whether, for example, one or two voiced bilabial consonants are present in *found both*, where this is not the case. However, the very fact that Gimson/Cruttenden see no need to comment on this difficulty suggests that this is a non-issue for them. Thus, although assimilation creates analytical problems for categorical phonological analyses of (t,d), the assimilated tokens again fit well into an integrated analysis of (t,d) as a CSP.

2.6 Coalescence

In the CSP literature, the term ‘coalescence’ is generally used to refer to the generation of, “a third ‘new’ segment (...) instead of two other abutting segments” (Nolan 1996:22). As with assimilation, there are examples of coalescence of both (t,d) consonants and their preceding consonants in the York data. All examples involve following /j/, as does Nolan’s example ([əʊzju] > [əʊzə] in *suppose you*). (72) and (73) illustrate /t#j/ sequences yielding [tʃ] and (74), taken from the same stretch of speech as (73), shows the preceding /s/ in *shortest* coalescing with following /j/ to yield a slightly lengthened [ʃː]. The latter two tokens are shown in Figure 10. (74) would presumably count as an instance of deletion in a variable rule analysis, whereas (73) would not, which seems to be imposing an artificial categorical divide on what looks like a continuum of phonetic explicitness.

- (72) *like [the baby] kept you up* [kʰɛpː tʃʊpː] *24 hours a night*
 (73) *the (.) longest you* [lɒŋgɪstʃɪ] *can wear is to there*
 (74) *the shortest you* [ʃaːtʰɪʃtɪ] *can wear is to there*

In (75) we observe a singleton word-final /z/ coalescing in the same way (but with additional devoicing):

- (75) *’cos you* [kʰəʃz] *can’t really do dances if you only get five turn up*

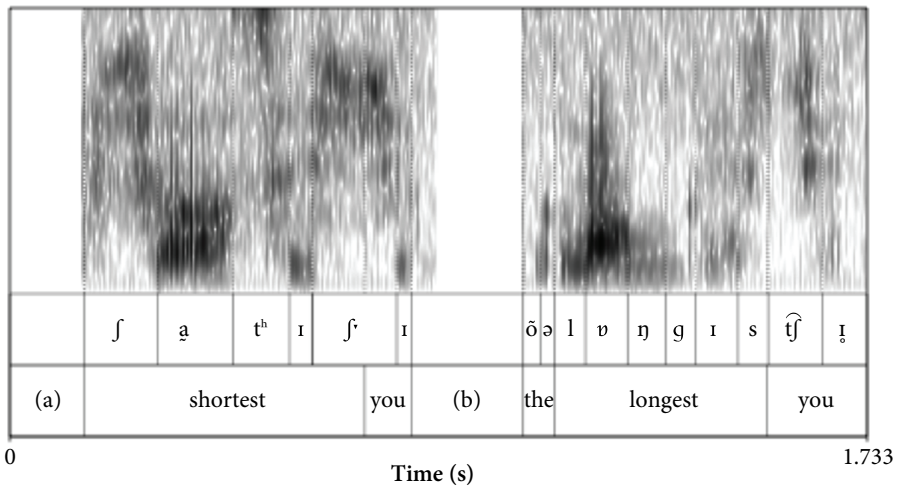


Figure 10. Spectrogram showing (in sequence) (a) *shortest you* (74) and (b) *longest you* (73); female speaker.

(76) shows a glottalised, nasalised glide resulting from the coalescence of properties from three segments, nasalisation from /n/, glottalisation from /t/ and labiality from /m/. It is illustrated in Figure 11.

(76) *and he didn't want me* [w̥w̥eɪ] *to leave*

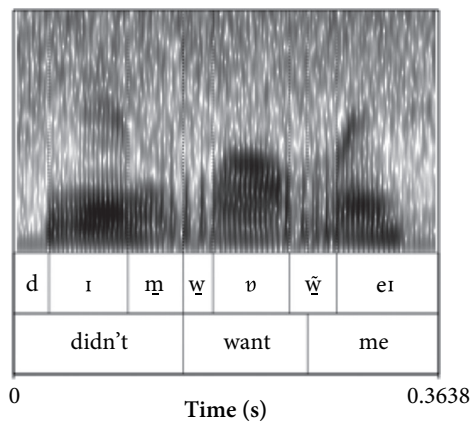


Figure 11. Spectrogram showing *didn't want me* (76); female speaker.

Coalescence in a more general sense is also seen in (t,d) between identical preceding and following consonants, where a single segment is generated from a sequence of two, with an intermediate ‘deleted’ (t,d) consonant. Sometimes these are more or less lengthened, as in (77) but frequently they are not, (78).

(77) *and just stabbed him* [n̠d͡ʒʊs:təbˈdəm]

(78) *it was my youngest son [juŋgɪsən] what caught me*

In (79) it is hard to decide whether the preceding /p/ is elided and the creaky voicing on the vowel is the reflex of the final /t/ (or indeed /pt/), or whether the /t/ is lenited or elided and there is a coalesced realisation of the preceding and following bilabial stops.

(79) *and he kept putting* [nikʰɛpʰʊtʰɪn] *it up and putting it up*

Tokens with following /ð/ are not part of the York (t,d) data because it constitutes a “neutralisation” context and such contexts are routinely excluded from analyses, but (80) is included here because the intervocalic [n] appears to be the result of progressive assimilation of nasality and stopping (unsurprising in /ð/-initial function words: see Manuel 1995), yielding what looks like a single, coalesced nasal:

(80) *if any of the schoolteachers found that* [faʊnəʔ] *you were misbehaving*

The issues raised for the analysis of (t,d) by tokens with coalesced preceding and following consonants (in both the narrow and broad sense) are essentially the same as those discussed under lenition and assimilation in §2.2 and §2.5 above, so we shall not revisit them here. Suffice to say that once again we find a range of examples of a well known CSP in both (t,d) and non-(t,d) contexts.

3. Discussion

In the light of the above detailed phonetic observations of the behaviour of (t,d) and other word-final consonants in York English, we now turn to the question of where they fit into a model of speech perception/production: do the facts about (t,d) merit its modelling as a variable phonological rule, as assumed in most of the variationist sociolinguistic literature? There are two aspects to the discussion, firstly whether (t,d) consonants are different from other word-final stops, which appears not to be the case, and then how the phenomena observed fit into the phonetics/phonology of English. Both, in my view, require if not resolution, then serious consideration before the further question of whether there is socioindexical variation in (t,d) and other word-final stops.

3.1 (t,d) and CSPs

The phonetic evidence surveyed in this paper has demonstrated that where direct comparison is possible word-final (t,d) consonants exhibit the same patterns of variability as other word-final stops, including variable pre-consonantal release characteristics and a range of degrees of lenition, crucially including full (auditory/acoustic) deletion. They also show parallel patterns of interaction with adjacent consonants resulting from Connected Speech Processes such as assimilation and cophonation. Such parallels have also been observed by Browman & Goldstein (e.g. 1990) using articulatory data from X-ray pellet-tracking: cluster-final /t/ and /d/ in *perfect memory* and *nabbed most*, when auditorily deleted, manifest a similar residual alveolar tongue gesture to word-final /n/ assimilating to following /p/ in *seven plus*. Moreover, even where direct comparisons across different consonants are not possible, there appears to be a plausible explanation in terms of CSPs for the whole range of variability observed in (t,d) clusters, including the behaviour of the first consonant of the cluster.

Furthermore, if (t,d) is a manifestation of general word-final CSPs, we would expect evidence of the cooccurrence of other CSPs in the surrounding speech. Thus (75) above shows voicing assimilation of the schwa of *you* to the preceding coalesced, devoiced [j] and following /k/, the type of cophonation Nolan focuses on (Nolan 1996:223). Cooccurrence of CSPs is illustrated more starkly in (76), where the whole sequence except the last word, *leave* shows decreased phonetic explicitness: the comments in §2.6 focussed on the coalescence at the end of *want* but in fact the whole sequence is highly reduced, *and he didn't want me to* being pronounced [əndʒiːmʷəŋwiːtə]. [d] and [j] are clearly articulated sequentially, but [n] and [d] are heterosyllabic, suggesting that the [d] is part of a coalesced pronunciation of *he di-*; the first [i] is nasalised in anticipatory assimilation to the following /n/, which assimilates in place to the /w/ of *want*; that /w/ is itself creaky-voiced, suggesting it bears a reflex of the final /t/ of *didn't*. In (81)²⁷ there is no acoustic or auditory evidence of any alveolar closure in the whole sequence /ntʃt/, close alveolar approximation not appearing until the following consonant /ð/. Note that nasality is also absent:

(81) *so they pinched the [piːʃðə] typewriter*

The (artificial) borderline between coalescence and cooccurrence breaks down at this point, but as noted at the outset, these categorisations are a descriptive convenience rather than a theoretical taxonomy. More importantly, the fact that

27. The (t,d) cluster in (81) would again be excluded from a variationist analysis because of the following 'neutralisation' context.

these patterns mirror general CSPs means that abstracting a specific (t,d) rule from examples such as (76) and (81) for the deletion of cluster-specific word-final /t/, rather than taking a holistic view of the sequence, would seem to call for independent justification.

Browman & Goldstein observe that the hitherto universally observed constraint ranking of following phonological segment on (t,d) is, “exactly what we would expect when we consider the consequences of gestural overlap” (Browman & Goldstein 1990: 367). The gestures best able to mask an alveolar closure gesture are precisely those which favour “deletion” of /t,d/, which leads them to conclude that, “the ordering of probabilities on deletion of final /t,d/ in clusters could follow directly from the view of deletion that we are proposing here, without these differential probabilities needing to be ‘learned’ as part of a rule” (Browman & Goldstein 1990: 368). Does a CSP analysis mean, then, that (t,d) should be viewed purely as a function of physical constraints which in turn vary as a function of factors such as speech rate? This is clearly not the case: there is plenty of evidence of dialect-specific patterning of the effect of following pause on deletion rates, for example (Tagliamonte & Temple 2005: 289), which must have a cognitive rather than a physical explanation. Individual speakers seem to show varying rates of “deletion”, so there must also be an idiosyncratic element in the phonetic implementation of word joins involving consonant sequences.²⁸ Speaker-specific manipulation of fine phonetic detail has long been known of and studied; for example, though physiological factors may play a role, in sex-specific variability, they cannot always explain the whole picture (e.g. Bladon et al. 1984; Temple 2000; see further Docherty 1992; Docherty & Foulkes 2005; Solé 2007). Indeed, in this volume Simpson shows how ejectives can be an epiphenomenon in one language and manipulated for interactional purposes in another.

Interactional effects are evident in the York (t,d) data too: as suggested by the contextualising comments accompanying some of the above examples, speakers appear to manipulate the phonetics of word-final stops for discourse purposes. Thus in (82) the speaker is recounting a sleepwalking episode after he had had rather too much to drink. His speech rate slows down and he produces a lengthened, preaspirated /s/ followed by a clear, but low-amplitude unaspirated released [t] followed by a pause lasting a second and *oh dear*. This is all clearly for comic effect, and the interviewer duly begins to giggle during the pause.

28. This has not been studied systematically in the York data. Impressionistically, speakers also appear to differ in the extent of phonetic explicitness in their speech overall. An empirical investigation of any correlation between lenition in word joins and other indicators of decreased explicitness would shed further light on this issue.

(82) *must have been completely lost (.)* [lv^hs:t] *oh dear*

(83) shows reported speech where the speaker is describing her rather large father threatening to take a “chopper” to the man who came round to means-test her for welfare payments in the 1930s. Again the utterance is intended to amuse and elicits the obviously anticipated laugh from the interviewer after the subsequent comment that, “the man never moved so fast in his life”.²⁹

(83) *“hand me [ˈʔandmiː] that so an’ so chopper”*

Note that the released [d] cooccurs with other indicators of fortition or increased phonetic explicitness such as the glottal stop at the beginning of *hand* (*h*-dropping is the norm for this speaker, including the *his* of the following sequence, which is produced vowel-initially) and the full vowel in *me*. Here too, then, the behaviour of (t,d) consonants appears to be consistent with surrounding CSPs rather than being independent of them.

Examples (82) and (83) and others like them suggest that, for these speakers at least, it is the surfacing of a released stop which is marked rather than its deletion. This said, however, in (84) a speaker who produces relatively high rates of surface cluster stops conversely twice elides the word-internal /d/ when quoting his friend’s girlfriend getting her own back after his nagging over her driving (leading to an embarrassing accident).

(84) *need the handbrake, [ambreˈkː] take the handbrake off [ambreˈkːvʃ], do this, do that*

3.2 Modelling variation in word-final stops

Does the cumulative evidence of speaker control over (t,d) mean that in fact (t,d) is a phonological rule after all, and the standard variationist account can be redeemed? In this view, the phonetic details observed in this paper would fall out from the production mechanism only after a variable categorical rule of deletion had applied. Such an argument is obviously a case of a *reductio ad absurdum*: the individual manipulation of fine phonetic detail first studied by phoneticians is now

29. Examples (82) and (83) show interactional effects in that they are intended to produce a response in the interlocutor. It seems likely that (t,d) and other word-final stops may also be manipulated interactively in the management of turn taking in the ways discussed by Simpson, this volume. As for Simpson, the nature of the data under discussion here make it difficult to be precise about this: sociolinguistic interviews are designed to elicit as much speech as possible from one party in the interaction, thus drastically reducing the number of potential and actual turn-transition points by comparison with naturally occurring conversation.

a generally accepted fact. The answer to the question of what (t,d) is and where (t,d) is properly to be located depends, then, on where the line is drawn in the grammar between phonology and phonetics, and how the interaction of cognitive and physical phonetic effects is modelled in the speech production model more broadly.

One possible answer to that question lies in the assigning of categorical processes to the phonology and gradient ones to the phonetic component of the grammar, and CSPs have played a central role in exploring this. Literature on categoricity vs. gradience in patterns of assimilation has been taken in the past to indicate that assimilation is either the result of a categorical phonological rule or of gradient phonetic constraints on articulation in fluent speech. Studies such as Ellis & Hardcastle (2002) show that in articulatory terms alveolar-to-velar assimilation may be gradient for some individuals and categorical for others, which is partly accounted for by accent differences. Aside from showing how the possibility of a total absence of the residual alveolar gesture is a problem for an Articulatory Phonology account of assimilation, they do not go into detail on the theoretical implications of their findings. However, such studies are taken by, for example, Bermúdez-Otero (2010b) to suggest that if (t,d) shows a mixture of gradient and categorical deletion then it must merit a two-step phonological derivation:³⁰

1. *phonology*: variable, categorical, morphologically sensitive
2. *phonetics*: variable, gradient, morphologically insensitive

(Bermúdez-Otero 2010b: 7)

The view of CSPs used as a framework for the present paper holds that they can be a function of both cognitive and physiological constraints, as Nolan notes with regard to assimilation: “it is a phenomenon over which speakers have control. This will provide further evidence that a greater amount of phonetic detail is specified in the speaker’s phonetic representation or phonetic plan than is often assumed” (Nolan 1992: 278; also cited by Ellis & Hardcastle 2002: 387). This implies a tripartite set of rules/constraints rather than a simple phonetics/phonology dichotomy, with the phonetic component consisting of both cognitive and physiologically constrained elements which can and do interact with each other.³¹ However, the potential existence of categorical deletion still need not necessarily entail that a categorical phonological rule is at work. Categorical deletion

30. It should be pointed out that for Bermúdez-Otero this is crucially also justified by the existence of the morphological constraint on (t,d) apparently found in many studies following Guy (1991).

31. This is not incompatible with Bermúdez-Otero’s position, which clearly includes gradient phonetic rules in the grammar and acknowledges the role of physiologically constrained processes in the production and perception of speech.

without a residual gesture may be viewed, as argued above, as a cognitively governed (phonetic) CSP at one end of a continuum of responses to the physiological challenge of producing an interconsonantal alveolar gesture. Thus at the ‘natural’ end speakers may be producing a full or partial alveolar gesture which is masked by surrounding gestures, whereas at the cognitive end of the scale they ‘choose’ not to. Indeed, Kühnert & Hoole (2004) report complex interactions of speaker-specific responses to articulatory challenges posed by alveolar-to-velar sequences showing the interaction of physiological and cognitive, language-specific and idiosyncratic effects. Cases with lenited alveolar gestures, which would be parallel to deleted (t,d) tokens, showed a range of qualitative differences between assimilatory and non-assimilatory contexts, showing that even in full “deletion” CSP effects are at work.

If some speakers can be shown by articulatory methods to be producing only categorical alternation between deletion and non-deletion with no gradient tokens, there may nevertheless be a case for saying that they represent a more advanced stage in a diachronic process of phonologisation of non-cognitive phonetic processes and their subsequent stabilisation as categorical phonological rules. This would follow the interpretation by Bermúdez-Otero & Trousdale (2011) of the inter-individual differences in assimilation patterns found by Ellis & Hardcastle (2002). However, so far as (t,d) is concerned there is no evidence in the literature for ongoing change: outside AAVE it does not show the sociolinguistic patterning (e.g. age-grading, a marked gender effect) which are expected to accompany change in progress, nor, to my knowledge, have published studies demonstrated real-time changes in patterns of deletion.³² In any case, the examples in the present paper of deletion of non-(t,d) consonants, could well also be categorical in the sense that a residual gesture could be entirely absent (e.g. (24) to (27), although we cannot say whether any of them produces nothing but categorical presence or absence). If categoricity is taken as requiring a phonological rule, then a phonological rule would also have to be formulated for these cases. Once again (t,d) does not look unique, and the problem of where to model these effects in the grammar remains.

As a phonetic-based approach, might Articulatory Phonology, which views phonological structure as, “an interaction of acoustic, articulatory, and other (e.g. psychological and/or purely linguistic) organizations” (Browman & Goldstein 1990: 341), provide a solution to the problem of situating (t,d) and related CSPs? (t,d) features prominently in early accounts of the theory, but there has been

32. Bybee (2002) implicitly assumes ongoing change in examining frequency effects on (t,d) in the context of lexical diffusion, but does not actually demonstrate that a diachronic process is underway.

very little articulatory study specifically of the variable since then. However, Lichtman's recent study of cluster and (mainly) non-cluster word-final /t/ examines data from the Wisconsin Microbeam Database and a complementary EMA study. While her results confirm the predictions of Articulatory Phonology regarding the effect of following phonological context on /t/ deletion, they also confirm Ellis & Hardcastle's finding that some individuals produce elided tokens without any residual alveolar gesture, which is not consistent with an AP account (Lichtman 2010; p.c.).

Interaction with other abstract levels such as morphophonology is another criterion which has been advanced for treating a phenomenon as phonological rather than phonetic (Tucker & Warner 2010: 318). The motivation for situating (t,d) in the (lexical) phonology originally was the apparent effect of morphology on its variability (e.g. Guy 1991). However, despite the many papers showing a statistical morphological effect, doubt has been cast by several recent studies on its veracity (see §1 above). Moreover, there is a fundamental methodological problem in the absence of large quantities of articulatory data: the evidence for the morphological constraint has generally been provided by auditory and acoustic data where it is impossible to tell whether the apparent deletion is categorical (and therefore by the logic of this account the result of a phonological, morphologically constrained rule) or gradient (and therefore the result of phonetic processes applying only after the morphological effect would have come into play). Lexical Phonology is a production-based model and so even a dual, 'rule scattered' account incorporating both categoricity and gradience stands on rather shaky ground in this respect until advances in articulatory sociophonetics allow us to collect large quantities of naturalistic conversational data, as acknowledged by Bermúdez-Otero (2010b).

The grammatical contrast between verbs with and without final *-ed* was invoked in pre-LP studies of (t,d) to account for the greater rates of retention of /t,d/ observed in past tense as opposed to monomorphemic forms. The role of contrastivity has received rather less attention in recent years than categoricity~gradience, but perhaps it would be fruitful to consider restricting an account of the phonology of (t,d) to stating their lexically contrastive terms, in which case both categorical and gradient deletion would be a phonetic phenomenon. A declarative, polysystemic analysis in the tradition of Firthian Prosodic Analysis (e.g. Robins 1970) would observe the limited distribution of word-final stops in general and of stop-final coda clusters other than (t,d) ones, and that there are very few minimal pairs contrasting cluster-final /t/ and /d/. Word-final postconsonantal stops would thus constitute a very restricted (sub?)system of phonematic contrasts. From the point of view of perception and comprehensibility, then, this view predicts that there is scope for a wide range of phonetic variability, which is indeed what we have observed in this paper.

To my knowledge no Firthian analyses yet exist of related English data. However, working from a very different perspective, Steriade (2000) explores critically the role of contrastivity in the categorisation of intervocalic flapping in American English (a phonetic phenomenon by this criterion). Her driving agenda is that, “the distinction between phonetic and phonological features is not conducive to progress and cannot be coherently enforced. It is unproductive because in order to understand phonological patterns one must be able to refer to the details of their physical implementation, in perception and production” (Steriade 2000: 314). Tucker & Warner explore the contrast between this view and the alternative strict separation of phonology and phonetics in the light of their analyses of the devoicing of nasals in Romanian. Having shown that the devoicing “derives from both phonetic and phonological causes” they point out that this does not necessarily entail the existence of two sharply delineated systems; it may simply be that, “all sound patterns fall somewhere on each of several dimensions that make up what we attempt to separate into phonetics and phonology” (Tucker & Warner 2010: 319). They argue that the answer to this is neither strict separation nor total integration but the classification of sound patterns on several, mostly continuous dimensions, “which all together make the phenomenon relatively phonetic or phonological” (Tucker & Warner 2010: 320). This approach would seem very promising for the analysis of word-final stops since it would obviate the need for a sharp dividing line between cognitively and physically constrained phonetic effects. We have seen the evidence for both here, and yet it is difficult to separate the two: as Kühnert & Hoole (2004) show, they interact at a highly detailed level, at least in assimilation, and their surface manifestations are often the same, and this seems also to hold for (t,d).

One aspect of the variable behaviour of final stops that these models do not cover, however, is the cooccurrence of CSPs. If it is the case that the pertinent dimension of sociophonetic variation is not the lenition/assimilation etc. of particular word-final segments, but the manipulation of phonetic explicitness across longer stretches of speech, a segmental based model would fail to capture the facts. Simpson demonstrates in this volume and elsewhere how restricting the analysis of variability to a single segment whose phonetics are governed by the immediate segmental context can obscure significant generalisations. In his analysis of glottals in Suffolk English (Simpson 1992), he examines the insight of Trudgill (1974) and Lodge (1984) that there are cooccurrence restrictions on glottalisation in some East Anglian varieties of English, and demonstrates that even one of the authors who drew attention to these misses some examples of the phenomenon because the analysis is couched in terms of derivational reduction rules which apply to individual segments. Simpson’s solution is inspired by the Firthian notion of “prosody”, a phonological construct which has phonetic exponents across a

given stretch (or “piece”) of speech. This would seem a promising avenue for exploration of the variability of word-final stops, although it should be noted that the Firthian approach is declarative, with strict separation between phonology and phonetics, and is therefore on the face of it not compatible with Steriade or Tucker & Walker’s advocacy of total or partial integration of the two.

4. Conclusions

This paper has, I hope, made a case for an answer to the first set of issues explored in the discussion, relating to the “what” of the title. It is clear from the data examined that the behaviour of word-final /t,d/ in clusters is not qualitatively different from that of other word-final consonants, either in their segment-specific physical manifestations or in their interactions with common Connected Speech Processes in this variety of English. This does not conclusively prove that a phonological analysis is wrong: CSPs could be part of the post-lexical phonetic implementation processes which interacts with the output of a variable phonological rule, as suggested by Bermúdez-Otero (2010a, b). However, since the CSP account seems perfectly adequate in accounting for the observed behaviour of word-final coronal clusters, it would seem that there is no need to invoke such a rule in the absence of positive evidence for an unambiguously phonological effect. It appears, then, that *WHAT* (t,d) is simply one manifestation of the general phenomenon that, in Browman & Goldstein’s words, “in casual speech (...) segments are routinely elided, inserted and substituted for one another” (Browman & Goldstein 1990: 359).

The other set of issues, that is *WHERE* this situates (t,d) and associated phenomena in the grammar, is less easy to resolve and depends on the place of this and other CSPs which, “can neither be modeled adequately at a symbolic, phonological level, nor left to be accounted for by the mechanics of the speech mechanism” (Nolan 1992: 280). But some well motivated model is needed in order to provide a sound basis for any sociophonetic/sociolinguistic analysis. The exploration in §3.2 of potential different frameworks for analysis was necessarily brief and far from conclusive, although it is clear that there are grounds for concluding that some are not satisfactory. With respect to more promising frameworks, the data reviewed here are insufficient to draw definitive conclusions about which approach to the phonology-phonetics interface best fits with the empirical observations of word-final stops. More articulatory data would be needed to implement a metric of gradience/categoricity,³³ for example, whereas a Firthian-inspired approach

33. Although even articulatory data would not be able to disambiguate all tokens, for example those with preceding /n/, which involve alveolar closure.

would require more data on other terms in the contrastive symbolic system and on the wider context. Ultimately the choice of model depends on the preference of the analyst, subject to the data. However, the exploration of some of these avenues with naturalistic data would provide opportunities for further advances in the interaction between sociophonetics and phonetic and phonological theory, and provide a better motivated model to serve as a foundation for the exploration of the social indexicality of these consonants.

References

- Barry, Martin C. 1985. "A palatographic study of connected speech processes". *Cambridge Papers in Phonetics and Experimental Linguistics* 4. 1–16.
- Bermúdez-Otero, Ricardo. 2010a. "Currently available data on English t/d-deletion fail to refute the classical modular feedforward architecture of phonology". *18th Manchester Phonology Meeting*, Manchester, 20th–22nd May 2010. <http://www.bermudez-otero.com/18mfm.pdf>
- Bermúdez-Otero, Ricardo. 2010b. "Morphologically conditioned phonetics? Not proven!" *On Linguistic Interfaces (OnLI) II*, Belfast, University of Ulster, 2nd December 2010. <http://www.bermudez-otero.com/belfast-handout.pdf>
- Bermúdez-Otero, Ricardo & Graeme Trousdale. 2011. "Pathways of phonological change". <http://ling.auf.net/lingBuzz/001235>
- Bladon, R. Anthony, Caroline Henton & Brian Pickering. 1984. "Towards an auditory theory of speaker normalization". *Language and Communication* 4. 59–69.
DOI: 10.1016/0271-5309(84)90019-3
- Broadbent, Judith M. 1999. "A new approach to the representation of coronal segments". *Issues in Phonology: Papers from an International Workshop* ed. by Stephen J. Hannahs & Mike Davenport, 1–25. Amsterdam & Philadelphia: John Benjamins.
- Browman, Catherine & Louis M. Goldstein. 1990. "Tiers in articulatory phonology, with some implications for casual speech". *Papers in Laboratory Phonology I. Between the Grammar and Physics of Speech* ed. by John Kingston & Mary E. Beckman, 341–376. Cambridge: Cambridge University Press.
- Buizza, Emanuela. 2010. "Frication and affrication of /t/ in RP English". *Language at the University of Essex (LangUE) 2010 Proceedings* ed. by Z. Absi, M. Althobaiti, V. Heyer, M. Ogawa, C. Patterson, & T. Soultatis, 16–28. <http://www.essex.ac.uk/langling/conferences/langue/>
- Buizza, Emanuela. 2011. "/t/ lenition in RP English spontaneous speech". *8th UK Conference on Language Variation and Change*, Edge Hill, 14th September 2011.
- Bybee, Joan L. 2002. "Word frequency and the context of use in the lexical diffusion of phonetically conditioned sound change". *Language Variation and Change* 14. 261–290.
DOI: 10.1017/S0954394502143018
- Coetzee, Andrees & Joe Pater. 2011. "The place of variation in phonological theory". *The Handbook of Phonological Theory* ed. by John A. Goldsmith, Jason Riggle & Alan Yu, 401–434. Oxford: Blackwell. DOI: 10.1002/9781444343069.ch13
- Crowdy, Steve. 1995. "The BNC spoken corpus". *Spoken English on Computer: Transcription, Mark-up and Applications* ed. by Greg Myers, Geoffrey N. Leech & Jenny Thomas, 224–235. Harlow: Longman.

- Cruttenden, Alan. 2008. *Gimson's Pronunciation of English*. London: Hodder.
- Docherty, Gerard J. 1992. *The Timing of Voicing in British English Obstruents*. Berlin: Foris Publications. DOI: 10.1515/9783110872637
- Docherty, Gerard J. & Paul Foulkes. 2005. "Glottal variants of /t/ in the Tyneside Variety of English: an acoustic profiling study". *A Figure of Speech* ed. by William J. Hardcastle & Janet Beck, 173–197. London: Lawrence Erlbaum.
- Ellis, Lucy & William J. Hardcastle. 2002. "Categorical and gradient properties of assimilation in alveolar to velar sequences: evidence from EPG and EMA data". *Journal of Phonetics* 30. 373–396. DOI: 10.1006/jpho.2001.0162
- Fabricsius, Anne. 2002. "Ongoing change in Modern RP: evidence for the disappearing stigma of t-glottaling". *English World-Wide* 23. 115–136. DOI: 10.1075/eww.23.1.06fab
- Foulkes, Paul & Gerard J. Docherty. 2005. "The social life of phonetics and phonology". *Journal of Phonetics* 34. 409–438. DOI: 10.1016/j.wocn.2005.08.002
- Guy, Gregory. 1991. "Explanation in variable phonology: an exponential model of morphological constraints". *Language Variation and Change* 3. 1–22. DOI: 10.1017/S0954394500000429
- Guy, Gregory & Charles Boberg. 1997. "Inherent variability and the obligatory contour principle". *Language Variation and Change* 9. 149–164. DOI: 10.1017/S095439450000185X
- Guy, Gregory & Sally Boyd. 1990. "The development of a morphological class". *Language Variation and Change* 2. 1–18. DOI: 10.1017/S0954394500000235
- Guy, Gregory, Jennifer Hay & Abby Walker. 2008. "Phonological, lexical and frequency factors in coronal stop deletion in early New Zealand English". *Laboratory Phonology 11 Conference*, Wellington, 30th June – 2nd July 2008. www.victoria.ac.nz/labphon11/LP11%20abstracts/Guy,%20Hay%20&%20Walker.pdf
- Hardcastle, William J. 1994. "Assimilation of alveolar nasals and stops in connected speech". *Studies in General and English Phonetics in Honour of Professor J. C. O'Connor* ed. by Jack Windsor-Lewis, 49–67. London: Routledge.
- Hayes, Bruce. 1992. "Comments on Chapter 10". *Papers in Laboratory Phonology II. Gesture, Segment, Prosody* ed. by Gerard J. Docherty & Robert Ladd, 280–286. Cambridge: Cambridge University Press.
- Hazen, Kirk. 2011. "Flying high above the social radar: coronal stop deletion in modern Appalachia". *Language Variation and Change* 23. 105–137. DOI: 10.1017/S0954394510000220
- Henderson, Jannette B. & Bruno H. Repp. 1981. "Is a stop consonant released when followed by another stop consonant?" *Haskins Status Reports in Speech Research* 67/68. 71–82.
- Kiparsky, Paul. 1985. "Some consequences of Lexical Phonology". *Phonology Yearbook* 2. 85–138. DOI: 10.1017/S0952675700000397
- Kühnert, Barbara & Phil Hoole. 2004. "Speaker-specific kinematic properties of alveolar reductions in English and German". *Clinical Linguistics and Phonetics* 18. 559–575. DOI: 10.1080/02699200420002268853
- Lichtman, Karen. 2010. "Testing Articulatory Phonology: variation in gestures for coda /t/". *Illinois Language and Linguistic Society Conference 2 (ILLS2)*, University of Illinois at Urbana-Champaign, 29th May 2010.
- Lodge, Ken R. 1984. *Studies in the Phonology of Colloquial English*. London: Croom Helm.
- Manuel, S. Y. (1995) "Speakers nasalize /ð/ after /n/, but listeners still hear /ð/". *Journal of Phonetics* 23. 453–476. DOI: 10.1006/jpho.1995.0033

- Myers, James. 1996. "The categorical and variable phonology of variable t-deletion in English". *International Workshop on Language Variation and Linguistic Theory*, University of Nijmegen, 10th September 1995 (revised version). <http://www.ccunix.ccu.edu.tw/~lngproc/t-deletion-manuscript.pdf>
- Nolan, Francis J. 1992. "The descriptive role of segments: evidence from assimilation". *Papers in Laboratory Phonology II. Gesture, Segment, Prosody* ed. by Gerard J. Docherty & Robert Ladd, 261–279. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511519918.011
- Nolan, Francis J. 1996. "Overview of English Connected Speech Processes". *Sound Patterns of Connected Speech. Description, Models and Explanation* ed. by Adrian P. Simpson & Matthias Pätzold, 15–26. Kiel, Institut für Phonetik und digitale Sprachverarbeitung.
- Roberts, Julie. 1997. "Acquisition of variable rules: a study of (-t,d) deletion". *Journal of Child Language* 24. 351–372. DOI: 10.1017/S0305000997003073
- Robins, Robert H. 1970. "Aspects of prosodic analysis". *Prosodic Analysis* ed. by Frank R. Palmer, 188–200. Oxford: Oxford University Press.
- Sankoff, David, Sali Tagliamonte & Eric Smith. 2011. *GoldVarb Lion. A multivariate analysis application*. <http://individual.utoronto.ca/tagliamonte/goldvarb.htm>
- Santa Ana, Otto. 1992. "Chicano English evidence for the exponential hypothesis: a variable rule pervades lexical phonology". *Language Variation and Change* 4. 275–288. DOI: 10.1017/S0954394500000818
- Schuppler, Barbara, Wim van Dommelen, Jacques Koreman & Mirjam Ernestus. 2009. "Word-final [t] deletion: an analysis on the segmental and sub-segmental level". *Proceedings of the 10th Annual Conference of the International Speech Communication Association (Interspeech 2009)*, Brighton, 6–10 September 2009, 2275–2278.
- Simpson, Adrian P. 1992. "Casual speech rules and what the phonology of connected speech might really be like". *Linguistics* 30. 535–548. DOI: 10.1515/ling.1992.30.3.535
- Smith, Jennifer, Mercedes Durham & Liane Fortune. 2009. "Universal and dialect-specific pathways of acquisition: caregivers, children and t/d deletion". *Language Variation and Change* 21. 69–95. DOI: 10.1017/S0954394509000039
- Solé, Maria-Josep. 2007. "Controlled and mechanical properties in speech. A review of the literature". *Experimental Approaches to Phonology* ed. by Maria-Josep Solé, Patrice Speeter Beddor & Manjari Ohala, 302–321. Oxford: Oxford University Press.
- Steriade, Donca. 2000. "Paradigm uniformity and the phonetics-phonology boundary". *Papers in Laboratory Phonology V. Acquisition and the Lexicon* ed. by Michael B. Broe & Janet Pierrehumbert, 313–334. Cambridge: Cambridge University Press.
- Stuart-Smith, Jane, Claire Timmins & Fiona Tweedie. 2007. "Talkin' Jockney: Accent change in Glaswegian". *Journal of Sociolinguistics* 11. 221–61. DOI: 10.1111/j.1467-9841.2007.00319.x
- Tagliamonte, Sali A. 1998. "Was/were variation across the generations: View from the city of York". *Language Variation and Change* 10. 153–91. DOI: 10.1017/S0954394500001277
- Tagliamonte, Sali A. & Rosalind A. M. Temple. 2005. "New perspectives on an ol' variable: (t,d) in British English". *Language Variation and Change* 17. 281–302. DOI: 10.1017/S0954394505050118
- Temple, Rosalind A. M. 2000. "Now and then: the evolution of male-female differences in the voicing of consonants in two varieties of French". *Leeds Working Papers in Linguistics and Phonetics* 8. 193–204.

- Temple, Rosalind A. M. Ms. "Rethinking (t,d): a new 'new look' at a variable construc[t]". [http://www.ling-phil.ox.ac.uk/files/research/Rethinking \(t,d\)2.pdf](http://www.ling-phil.ox.ac.uk/files/research/Rethinking_(t,d)2.pdf)
- Trudgill, Peter. 1974. *The Social Differentiation of English in Norwich*. Cambridge: Cambridge University Press.
- Tucker, Benjamin V. & Natasha Warner. 2010. "What it means to be phonetic or phonological: the case of Rumanian nasals". *Phonology* 27. 289–324. DOI: 10.1017/S0952675710000138
- Wells, John C. 1982. *Accents of English*. Cambridge: Cambridge University Press.
- Wolfram, Walt. 1969. *A Sociolinguistic Description of Detroit Negro Speech*. Washington, D.C.: Center for Applied Linguistics.
- Wolfram, Walt. 1993. "Identifying and interpreting variables". *American Dialect Research* ed. by Dennis Preston, 193–221. Amsterdam & Philadelphia: John Benjamins.

New parameters for the sociophonetic indexes

Evidence from the Tuscan varieties of Italian

Giovanna Marotta

Università di Pisa

A sociophonetic analysis of the main phonological processes occurring in Tuscan Italian is presented within a global proposal of a new, original set of parameters of variation. After a general discussion on the sociophonetic indexes and the illustration of phonological processes occurring in the local pronunciation of Italian, the parameters of the new model are metaphorically identified as properties of solids, i.e. shape, size, thickness and weight. In the last section of the paper, the sociophonetic parameters proposed are compared with the socially-marked variables proposed by Labov (2001), showing analogies and differences. The advantages derivable from the model proposed are finally discussed, with the explicit acknowledgement of the need for the inspection of the phonological system in sociophonetic analysis.

1. Introduction

Sociophonetics has a long history and a brief history at the same time. In Western culture, with reference to language variation in terms of difference in education and diatopic characterization, it goes back to at least the Republican period of Roman culture, when Catullus was joking with Arrius in *Carmen* 84:

Chommoda dicebat, si quando commoda vellet
dicere, et insidias Arrius hinsidias,
(...)
Ionios fluctus, postquam illuc Arrius isset,
iam non Ionios esse sed Hionios.

The example is not casual, as the typical lenition which Catullus was probably referring to in his *Carmen* will be considered, with respect to its modern Tuscan counterpart, in the following pages.

A great deal of modern sociolinguistic research is essentially sociophonetic, since the phonetic level of language is traditionally the most investigated in this domain. If we consider traditional dialectology too, it is easy to observe that the phonetic analyses are much more numerous than the morphological or syntactic ones, especially with reference to the Italian domain.¹ On the other hand, the term 'sociophonetics' is quite recent in the scientific literature of linguistics.² It is probably not a chance the fact that in the index of both volumes of *Principles of Linguistic Change* (Labov 1994, 2001), the term sociophonetics is missing, and its field is mostly covered by 'sociolinguistics' and 'sociolinguistic patterns'.³

In short, it is quite clear that a sociophonetic perspective entails the study of those phonetic variations in speech that are socially driven. However, as normally acknowledged, the correlation between speech variation and social structure is the basic tenet of traditional sociolinguistics since the beginning of its history, with the seminal work by Labov in the mid-1960s.

What seems to be the peculiar, and maybe innovative, feature of sociophonetic research is probably the wish of joining sociolinguistics with experimental phonetics.⁴ However, as Labov himself (2006: 500) expressly recognizes, in the sociolinguistic enterprise phonetic acoustic analysis played a major role since the earlier studies of sound changes in progress occurring in many speech communities. Being specifically focused on the interaction between the social context and phonetic controlled experimentation, sociophonetics appears to be more a sub-field of sociolinguistics than an autonomous discipline. Its tenets, both theoretical and methodological, are borrowed from sociolinguistics on one hand, and from experimental phonetics, on the other.

There is no doubt that phonetic variability may index social meaning: the large amount of data coming from different areas of the world has been witnessing this

1. As a matter of fact, the situation concerning the Italian dialects requires a special attention, because the amount of data collected throughout the last centuries is indeed very rich and remarkable, both on the quantitative and qualitative sides.

2. As is well-known, this term goes back to a PhD thesis discussed in 1974 at UCL (Deshaies-Lafontaine 1974). Moreover, as Foulkes & Docherty (2006) observe, it has neither been used in a consistent way nor often adopted till present time.

3. However, it is interesting to observe that some years later, the same Labov (2006) wrote a very stimulating article in the special issue of the *Journal of Phonetics* devoted to *Modelling sociophonetic variation*, where he discusses some of the typical theoretical and methodological problems connected with a sociophonetic framework, and claims that "sociophonetic studies are not disjointed from the broader field of sociolinguistics" (Labov 2006: 501).

4. As Jannedy & Hay (2006: 406) wrote, sociophonetic researchers "feel they straddle the divide between sociolinguistics and phonetics".

fact for a long time. Nevertheless, the way in which the speaker/listener of a linguistic community can arrange the huge variability normally occurring in speech is still not clear. Neither is it clear whether and how sociophonetic variation informs the cognitive patterns of the individuals. In other words, although human utterances normally give information about social factors, we may ask whether it makes any sense to describe all such sociophonetic indexing. As correctly Labov (2006: 508) observes, “no matter how narrowly we code variation, there still remains a residue of free variation”. And free variation has no cognitive value.

On reverse, we would like to assume a more abstract, cognitive-oriented point of view. Our aim is that of presenting the sociophonetic analysis of a set of processes occurring in Tuscan Italian with crucial reference to the phonological system, i.e. to the abstract representation of sounds and sound categories. The major goal of this paper is to propose a new, original set of parameters for the analysis of sociophonetic indexes, based on the metaphor of speech as a solid body.

The paper is organized as follows. In §2, a preliminary discussion of the nature of sociophonetic indexes is presented, in relation to some complex issues which are only partially debated in the existing literature. In §3, the linguistic repertoire of Tuscany is briefly introduced with special reference to the main phonological processes occurring in the local pronunciation of Italian. In the subsequent paragraphs (§4 through §8), a new model for sociophonetic analysis is proposed. Being the model grounded on the metaphor of solids, its parameters are metaphorically identified and defined as the properties of solid bodies, i.e. shape, size, thickness and weight. Then, the possible interactions among the parameters is shortly addressed (§9). In the last sections of the paper (§10–§11), the sociophonetic parameters proposed are compared with the socially-marked variables proposed by Labov (2001), showing analogies and differences. The advantages which can be derived from the model are finally discussed, with the explicit acknowledgement of the need for the inspection of phonological systems in sociophonetic analysis.

2. The sociophonetic indexes

Relevant information about the speaker as a member of a specific linguistic community normally derives from sociophonetic indexes, which are often grounded in the so-called fine phonetic details (Local 2003; Hawkins 2003, 2010; Carlson & Hawkins 2007).⁵ However, not all fine phonetic details which can be found in

5. As is well-known, the term ‘index’ is basically used with reference to Peirce’s semiotic theory since the beginning of last century: an index refers to an individual object independently of any resemblance to it, i.e., it is a sign that denotes its object by virtue of an actual connection

the acoustical signal have necessarily a corresponding perceptive value for the listener, although the sociophonetic indexes are necessarily carried on by some phonetic substance.

Sociophonetic indexes are normally conceived of as gradient rather than categorical. In sociolinguistic research notions like *continuum* and gradient scale are traditionally employed instead of category and discreteness: e.g., idiolects and linguistic varieties merge one into the others without a clear definable boundary; the linguistic phenomena sensitive to the socio-cultural factors show a gradual nature more than a binary one; phonemes cannot be conceived any more as granitic entities, identical in every situation and in all the speakers of a language or a dialect (see for instance Labov 1972, 1994; Thomason & Kaufman 1992; Coulmas 1998, 2001).

However, in our opinion, sociophonetics cannot entirely dispense with a sort of discrete representation of sounds, if its goal is to capture the linguistic competence of the speakers (see below, §9 and §10). The perception of speakers has to be grounded on some kind of categorical units, in the sense of cognitive entities located in the mind of the speakers/listeners of any linguistic community.

Indexicality is a double-face entity: on the one hand, it can be considered as the measure of correlations between phonetic variable forms and social factors; on the other one, it may also refer to the awareness of these correlations by speakers. More or less, this difference covers the distinction made by Silverstein (2003) between First Order Indexicality and Second Order Indexicality: in First Order Indexicality, reference to patterns of correlation between specific linguistic forms and social factors is meant, whereas in Second Order Indexicality the awareness of the sociophonetic correlations by the speakers and listeners of a given social community is involved. Moreover, the awareness can be overt or tacit, according to the degree of prestige of the linguistic variation (Trudgill 1972; Labov 1966, 1975, 2001).

As a matter of fact, we believe that a deep inspection into the nature of sociophonetic indexes is needed in order to catch their crucial aspects. First of all, the phonetic features encompassing sociolinguistic variation do not have the same status, both in the speaker's awareness and with reference to the linguistic system; therefore, they cannot be considered in the same way. Their status can be different, because:

involving the two entities (sign and object); therefore, there is a causal relation between *signans* and *signatum* (Peirce 1972, 1991). Let us quote two relevant passages with specific reference to this notion: "an *index* is a sign which would, at once, lose the character which makes it a sign if its object were removed, but would not lose that character if there no interpretant. Such, for instance, is a piece of mould with a bullet-hole in it as sign of a shot" (Peirce 1991: 239); "an *index* is a sign really and in its individual existence connected with the individual object" (Peirce 1991: 251).

- a. the perception of the sociophonetic features by the speakers can vary: for some features, there can be no perception at all, whereas for others the perception can be very fine-grained. In other words, the degree of awareness is variable, and can even be completely absent; in Silvestein's terms, Second Order Indexicality is not always present in sociophonetic indexes;
- b. sociophonetic features can be more or less central in the phonological system of a language. A marginal or central position occupied by the variation connected with a specific feature has different effects on usage as well as on language structure.

With regards to the last point, we believe that sociophonetic analysis cannot entirely dispense with phonology, inasmuch phonetic variation occurring in a community of speakers can adequately be described and interpreted only with crucial reference to a shared phonological system. To stress the relation between phonetic variation and the abstract level of phonological representation is unusual within the framework of sociophonetics, which is normally focused on the behavior of single speakers. On reverse, we would like to assume a system- more than a speaker-oriented theoretical and methodological point of view.

Our proposal of new parameters in the analysis of sociophonetic indexes is grounded on the acknowledgment that the variable forms in speech do not have: (i) the same socio-cultural status; (ii) the same degree of speakers' awareness; (iii) the same domain, or scope.

Not always these three aspects are dealt with in sociophonetic analyses. Sometimes the speakers' awareness is missing, some others the scope is ignored; very rarely, the phonological structure on the background of the phonetic variation is considered.

Our focus here is on sociophonetic variation occurring in the Tuscan region. The typical phonological processes occurring in Tuscan Italian have already been described in the literature produced within the traditional stream of dialectology, which has a long-standing and distinguished tradition of studies in Italy (for instance, see Giannelli 1976). Here, the main Tuscan processes will be discussed with reference to a new model of sociophonetic variation which should allow us to better enlighten their properties as well as their position within the phonological structure.

In particular, we would like to propose that phonetic cues expressing allophonic variation can be described and classified according to a model referring to the properties of the solids, such as shape, size, thickness and weight. These properties can be very useful in the interpretation of the empirical phenomena as well as in the recognition of their constraints, both distributional and socio-cultural. Hopefully, they could also give relevant cues for a better explanation of linguistic change.

A central feature of any sociolinguistic approach is the acknowledgement of the intrinsically dynamic nature of language: far from being a static monolith, language is considered as a living organism, it is viewed in its actual usage, as a social medium of communication among people. If we acknowledge that sociophonetic variables can be assigned different values as regards their status, scope and awareness degree, we will be able to better represent the intrinsically dynamic nature of language structure.

As already mentioned, the point of view assumed here is more system-oriented than speaker-oriented. In other words, we are not concerned with sociophonetic indexes in strict relation with different or special groups of speakers, or particular styles of speaking. Neither will we present any fine-grained acoustical analysis of the relevant Tuscan phonological processes, because all these phenomena have already been analyzed in detail by Italian scholars in the previous literature. We rather focus on the general properties of sociophonetic variation, in order to show how they may interact with each other and whether they can affect the linguistic structure.

In short, the present article would like to be a modest contribution to adequately handling sociophonetic variability occurring in Italian varieties, with the hope that other scholars could find this proposal useful in future sociophonetic research.

3. The linguistic repertoire of Tuscany

As is well known, Standard Italian is based on the Tuscan dialect; more precisely, on Florentine as spoken in the late 13th/14th centuries.⁶ The vowel system is the same in Standard Italian and in the Tuscan variety of Italian: seven phonemes (/a ε e i ɔ o u/) in stressed syllable, five in unstressed position (where the opposition between mid-open and mid-closed vowels is not effective). No length contrast occurs in vowels, although a phonetic rule of vowel lengthening applies in the context of stressed open non final syllable (e.g. *pane* ['pa:ne] 'bread', *piede* ['pje:de] 'foot', *tavolo* ['ta:volɔ] 'table' vs. *pasta* ['pasta] 'pasta', *mangiò* [man'dʒɔ] '(s/he) ate'). As for the consonant system, the Italian repertoire, as well as the Tuscan one, is rather simple: the places of articulation for obstruents are bilabial (/p b/), labiodentals (/f v/), dental (/t d s/), palatal (/ʃ/) and velar (/k g/). The sonorants are /r l ʎ/, the nasals /m n ɲ/, with the allophones [m̥ n̥]. The picture is enriched by the occurrence of a set of affricates (/ts dz tʃ dʒ/) and by the feature of length, which applies to almost all consonants of Italian and Tuscan as well. As a matter of fact, gemination is certainly one of the most relevant phonological features of Italian consonant system, even in comparison

6. The Florentine dialect acquired its prestige because of the masterpieces of Italian literature written by Florentine authors, in particular Dante and Boccaccio (cf. Lepschy and Lepschy 1977). For a general picture of the phonology of Italian, see Bertinetto & Loporcaro (2005).

with the other Romance languages: the contrast between C and C: exhibits a very high functional load in most varieties. Nevertheless, Northern dialects lack long consonants, and for this reason Northern speakers of Italian may show a less consistent and systematic correlation of gemination.

Although the phoneme inventory of Standard Italian is basically the Florentine one, a great amount of phenomena occurring in the Tuscan varieties of Italian are absent in the standard national language. Our focus will be on the Tuscan varieties of Italian more than on the dialects spoken in Tuscany, the so-called *vernacoli*. Despite the closeness between the standard language and the dialects spoken in Tuscany, it is indeed possible to single out the structural properties of Tuscan varieties from the general ones belonging to Italian (see for instance Agostiniani & Giannelli 1990).

The main phonological processes typical of the Tuscan pronunciation of Italian can be summarized as follows (see Giannelli 1976, 1988; Giannelli & Savoia 1978, 1979–80; Agostiniani 1989, 1992; Marotta 1995, 2008):

1. the so-called *gorgia toscana*, that is the typical lenition of stops in intervocalic position; e.g. SI *pipa* 'pipe', Tsc ['pi:ɸa], SI *vita* 'life', Tsc ['vi:θa], SI *poco* 'little', Tsc ['pɔ:xo];⁷
2. the spirantization of palatal affricates in intervocalic position; e.g. SI *amici* 'friends' [a'mirtʃi], Tsc [a'mi:ʃi], SI *bicicletta* 'bike' [bitʃi'kletta], Tsc [biʃi'xletta], SI *agile* 'agile' [a:dʒile], Tsc [a:ʒile];
3. the affrication of /s/ after alveolar sonorants; e.g. SI *salsa* 'sauce', Tsc ['saltsa], SI *borsa* 'bag', Tsc ['bortsa], SI *pensiero* 'thought' Tsc [pen'tsje:ro];
4. the apocope, that is the deletion of postvocalic vowels in word-final position; e.g. SI *la mia mamma* 'my mother', Tsc *la mi' mamma*; SI *la sua sorella* 'your sister', Tsc *la su' sorella*; SI *dei bambini* 'some children', Tsc *de' bimbi*;
5. the sandhi process called *rafforzamento* (or *raddoppiamento fonosintattico*), a consonant lengthening process taking place at word boundary after some function words or after an oxytone word; e.g. *io e Maria* [e 'm:aria] 'Mary and I', *da Roma* [da 'r:oma] 'from Rome', *un tè leggero* [tɛ lɛ'd:ʒɛ:ro] 'a light tea', *parlò forte* [par.lɔ 'forte] '(s/he) spoke aloud';⁸
6. the truncation of infinitive verbal forms; e.g. SI *mangiare* 'to eat', Tsc [man'dʒa], *vedere* 'to see', Tsc [ve'de], SI *sentire* 'to hear', Tsc [sen'ti].⁹

7. Here and in the rest of the paper, SI = Standard Italian, Tsc = Tuscan.

8. The historical source of this process is the assimilation at word boundary between an etymological final consonant and a following initial consonant (see Loporcaro 1997). Central and Southern dialects as well as their respective varieties of Italian share the process, although the lexical distribution can vary from place to place; the phenomenon is lacking in the North of Italy, where no length contrast occurs, at least in the dialects.

9. This last process has a morphophonological nature, since it applies to a specific category of forms in the Verbal Inflection; see Marotta (2000) for details.

Other phonological processes occurring in Tuscany, such as the rhotacism of the lateral in pre-consonantal position, the palatalization of the lateral consonant, the vowel raising and lowering in pre-tonic and post-tonic position (cf. Giannelli 1976), have a more marked vernacular status, belonging to the dialects and not to the local pronunciation of Italian. Therefore, they will not be taken into account in this analysis.

On the other hand, a special process occurring in the areas of Pisa and Leghorn (North-West part of the region) has to be mentioned here, since it exhibits a quite relevant sociophonetic status, with reference to speaker identity as well as socio-cultural status. This process is the velarization of the lateral consonant when geminated or in syllable coda position; e.g. *bello* 'nice' ['beʎ:o], *alto* 'high' ['aʎto] (Marotta & Nocchi 2001). The phenomenon has been only cursorily studied in the previous literature on Tuscan varieties and dialects. Therefore, considering both its probable sociophonetic status and its relative newness, we have included the analysis of *l*-velarization in the present review, despite its marginal geographical position.

In Figure 1, a map of Tuscany is presented, with the indication of the ten regional districts. As is well-known, the varieties spoken in the Eastern area (basically, the district of Arezzo) and in the North-Western area of the region (the district of Massa-Carrara) do not show the typical phonological features of the Tuscan varieties, whereas they share phonological and morphological properties with Central and Northern varieties of Italian, respectively (Giannelli 1976).

4. The model and its parameters

Before analyzing the main sociophonetic indexes characterizing the Tuscan varieties, the descriptive parameters which will be used have to be identified and defined. They will be discussed in terms of a metaphor, the metaphor of solids. Sociophonetic variation can indeed be viewed as a solid body, i.e. an entity occupying a specific space in the domain of language and occurring in a delimited time.

As cognitive linguistics has been showing for many years, our metaphors always start from a concrete concept, normally grounded on our sensorial and perceptive experience. In this case, the solid metaphor allows us to conceive sociolinguistic variation in a more concrete way, to 'see' the linguistic properties as physical properties of objects having a physical nature. Being grounded in a conceptual metaphor, the parameters assigned to sociophonetic indexes have to be considered as descriptive more than theoretical and formal.

In the geometry of solids, starting from Euclid in ancient Greece, a solid is defined as a geometrical entity with three basic dimensions: *x*, *y* and *z*. It represents



Figure 1. Map of Tuscany, with indication of the ten regional districts.

a section of space defined by its surface. In geometry, two major features of solid bodies are recognized: *SHAPE* and *VOLUME* (Albert 1949). Solids have different *SHAPE*: a cube has a specific form, which is different from that of a cylinder or a parallelepiped. A solid has a *VOLUME*, or *SIZE*: for instance, a cube or a cylinder may be small or big, and smaller or bigger than other solids.

Furthermore, when a solid is put in a physical environment, it is given two other relevant properties: *WEIGHT* and *THICKNESS*. With reference to the force of gravity, in the real world, a solid has a *WEIGHT*; in relation with its front view, a solid has a *THICKNESS* too.

Can the quoted properties of solids be used as metaphorical parameters for the interpretation of phonological processes? Can these parameters be referred to the phonetic features implementing sociolinguistic variation? This is our present challenge. In the next paragraphs, the quoted properties of solids will be considered as metaphorical parameters of sociophonetic variation and will be discussed in an analytical way, showing at the end whether and how they can interact with each other.

Before presenting our proposal, a synthetic list of the phonological processes which will be dealt with is given. According to what has been said in §3, they are:

1. *gorgia toscana*;
2. de-affrication of palatal affricates;
3. apocope;
4. *rafforzamento fonosintattico*;
5. *s*-affrication;
6. *l*-velarization (Pisa and Leghorn);
7. oxytone infinitives.

Each process will be discussed in relation to its distributional context and making use of the sociophonetic parameters proposed.

5. *SHAPE*

The *SHAPE* of a sociophonetic variation is simply represented by the description of a given phonological alternation: the set of linguistic features and their context of occurrence gives the process its *SHAPE*. The phonetic indexes expressed by the speakers' behavior describe the *SHAPE* of a sociophonetic phenomenon, inasmuch they not only exhibit specific phonetic and prosodic details, but also make reference to the phrasal context and to the interface with the other levels of the grammar. Examples of *SHAPE* of the Tuscan sociophonetic phenomena listed above are presented in this section.

5.1 *Gorgia toscana*

Gorgia toscana has the *SHAPE* of an articulatory weakening of plosive consonants, both voiceless and voiced, in postvocalic context, not only within the word domain, but also in the phrase domain.¹⁰ Some examples of the phenomenon are given in (1):

- (1) SI *fico* 'fig', Tsc ['fi:ho]; SI *lato* 'side', Tsc ['la:θo]; *pipa* 'pipe', Tsc ['pi:ɸa]; SI *piega* 'fold' ['pje:ya]; SI *la casa* 'the house', Tsc [la 'ha:sa]; SI *la torta* 'the cake', Tsc [la 'θorta]; SI *la pasta* 'the pasta', Tsc [la 'ɸasta].

10. The literature on the Tuscan *gorgia* is very rich. For a general survey of the topic, the reader is referred to Giannelli (1976), Giannelli & Savoia (1978, 1979–80, 1991), Agostiniani & Giannelli (1990), Marotta (1995), Giannelli & Pacini (1998). More recently, the phonetic and acoustical aspects of this phonological process have been investigated by Marotta (2001/2004, 2008).

In formal terms, using a rule-based framework, the *gorgia* has the following *SHAPE*:

$$(2) \quad C [-\text{cont}] \rightarrow [+ \text{cont}] \quad / \quad V (\#) _ (L,G) V$$

In an autosegmental representation, its *SHAPE* is:

$$(3) \quad \begin{array}{ccccc} [-\text{cont}] & \rightarrow & [+ \text{cont}] \\ | & & | \\ V & C & V & & C \\ | & | & | & & | \\ x & x & x & & x \\ | & | & | & & | \\ N & O & N & & O \end{array}$$

5.2 De-affrication of palatal affricates

A similar consonant weakening, i.e. the de-affrication of palatal affricates (/tʃ, dʒ/ > [j, ʒ]) normally taking place in Tuscany can be considered as a special case of *gorgia*, inasmuch as it is a weakening process occurring in the same postvocalic context, within the word domain as well as in the phrase domain (Marotta 1995, 2008). Some examples are given in (4):

- (4) SI *amici* ‘friends’ [a'mi:tʃi], Tsc [a'mi:ʃi]; SI *bicicletta* ‘bike’ [bitʃi'klet:a], Tsc [biʃi'hlet:a]; SI *agile* ‘agile’ [a:dʒile], Tsc [a:ʒile]; SI *la ciabatta* ‘the slipper’ [la tʃa'bat:a], Tsc [la fa'βat:a], *la giacca* ‘the jacket’ [la 'dʒak:a], Tsc [la 'zak:a].

In autosegmental terms, the process shows the following *SHAPE*:

$$(5) \quad \begin{array}{ccccccc} C [+ \text{cor} [-\text{front}]] & \rightarrow & C & / & V (\#) _ (L,G) V \\ & / & \backslash & & | \\ [-\text{cont}] & [+ \text{cont}] & & & [+ \text{cont}] \\ & \backslash & / & & | \\ V & C & V & & C \\ | & | & | & & | \\ N & O & N & & O \end{array}$$

5.3 Apocope

Another typical process of the Tuscan dialects as well as of the Tuscan pronunciation of Italian is apocope, i.e. the deletion of the unstressed final vowel in context of hiatus, i.e. in postvocalic position (Marotta 1995). The deleted vowel is normally the high front vowel /i/, which carries an important morpho-phonological

function as a marker of gender and number in the case of noun declension and of person and number in the case of verb conjugation. However, in the noun phrases, other vocalic segments can be deleted too (-a, -o, -e), as results from the examples in (5):

- (6) SI *dei bambini* 'some children', Tsc *de' bimbi*; SI *mangiai molto* '(I) ate a lot', Tsc *mangia' molto*; SI *poi vengo* 'I come after', Tsc *po' vengo*; SI *il tuo fratello* 'your brother', Tsc *il tu' fratello*; SI *la sua sorella* 'his/her sister', Tsc *la su' sorella*; SI *le sue bimbe* 'his/her girls', Tsc *le su' bimbe*.

In formal terms, using a rule-based framework, apocope will have the following *SHAPE*:

- (7) V [-stress] → [Ø] / V _ #

The only constraints applying on this kind of vowel deletion pertain to the prosodic structure, since the target of the process has to be unstressed and preceded by a vocalic nucleus.

It is worthwhile to underline the fact that apocope interacts with *gorgia* inasmuch it feeds the contexts triggering the spirantization of the stop consonants, see examples in (8):

- (8) SI *arrivai tardi* '(I) came late', Tsc [arri:va 'θardi]; SI *hai capito?* 'Have you understood?', Tsc [a xa'φi:θo]; SI *hai preso il suo cappello* '(you) took his/her hat', Tsc [a 'φre:so il,su xa'p:el:o].

In other words, vowel deletion does not block *gorgia* because the output of the process is still a context favoring consonant spirantization.

5.4 Rafforzamento fonosintattico

A typical feature of Tuscan dialects as well as of Tuscan Italian is the sandhi process called *rafforzamento fonosintattico* (henceforth, RF). This consonant strengthening process takes place at word boundary after some functional words or after a final stressed final vowel (Agostiniani 1992; Loporcaro 1997; Borrelli 2002). It consists in the gemination of the initial consonant of the following word; e.g.:

- (9) *un tè* [fɪ]orte 'a strong tea'; *mangiò* [tɪ:]utto '(s/he) ate everything'; *io e* [lɪ:]ui 'he and I'; *da* [m:]ilano 'from Milan'; *una città* [pɪ:]ulita 'a clean city'.

Central and Southern dialects as well as their respective varieties of Italian share RF with the Tuscan varieties, although in non Tuscan varieties the distribution of the process varies from place to place (Loporcaro 1997; Fanciullo 1997) and

appears to be more morphologically-driven. By contrast, the phenomenon is completely missing in the North of Italy, where there is no consonant length contrast.

In terms of a phonological rule, a possible *SHAPE* of the process is:

$$(10) \quad C \rightarrow C: / \begin{matrix} [V] \\ [+stress] \end{matrix} \# _ \left\{ \begin{matrix} [C, +snrt] \\ G \end{matrix} \right\}$$

In Tuscany, the process is sensitive to the prosodic context too, since the target consonant is particularly lengthened in the case of a stress clash, that is when the first word ends with a stressed vowel and the following word begins with a stressed syllable: consider for instance *città grande* ‘a big city’ [tʃi:ta 'g::rande] vs. *città grandissima* ‘a very big city’ [tʃi:ta g:ran'dis:ima]; *caffè nero* ‘black coffee’ [ka:fɛ 'n::e:ro] vs. *caffè nerissimo* ‘very black coffee’ [ka:fɛ n:e'ris:imo] (Marotta 1983–1986). This over-lengthening of the consonant in case of stress clash is the effect of a prosodic constraint working at the surface phonetic level and showing the usual aspects of gradualness typical of such a level of analysis (see also §5.6).

5.5 s-affrication

A special attention has to be devoted to the process of /s/-affrication occurring in post-consonantal context; in particular, in Tuscany the process is triggered by a preceding alveolar sonorant (Celata 2008, 2009), as shown in (11):

- (11) SI *salsa* ‘sauce’, Tsc ['saltsa]; SI *borsa* ‘bag’, Tsc ['bortsa]; SI *pensiero* ‘thought’, Tsc [pen'tsje:ro]; SI *ansioso* ‘anxious’, Tsc [an'tsjo:so]; SI *orso* ‘bear’, Tsc ['ortso]; SI *arso* ‘burned’, Tsc ['artso]; *morso* ‘bite’ Tsc ['mɔrtso].

The process applies not only within the word domain, but also in the phrase domain; e.g. *il sole* ‘the sun’, Tsc [il 'tso:le]; *del sale* ‘some salt’, Tsc [del 'tsa:le]; *in sala* ‘in the dinner room’, Tsc [in 'tsa:la]; *con Simone* ‘with Simon’, Tsc [kon tsi'mo:ne], or even [ko tsi'mo:ne], with total assimilation of the nasal consonant in coda position.

A possible formalization of this process is:

$$(12) \quad [s] \rightarrow [ts] / C [+snrt] _$$

This strengthening process is normally not considered a typical stereotype of Tuscan speech, as it is spreading through a large area of Central and Southern Italy (Holtus et al. 1988; Telmon 1996). Tuscan speakers do not seem to have explicit awareness of the process. However, this process of allophonic variation constitutes a possible source of phonemic ambiguity, since the affrication of a

sibilant gives rise to an affricate and Italian has a dental voiceless affricate as a phoneme, which can occur even in the context triggering /s/ affrication (e.g., *alza* ‘(he) raises’; *danza* ‘danse’).

Therefore, a ‘near-merger’ of the two phonological categories arises,¹¹ because when a Tuscan speaker-listener hears a phonetic string like [-nts-], [-rts-], [-lts-], he/she cannot know whether there is a /s/ or a /ts/ phonemic category, unless he knows the meaning of the word and how it is written. The process of s-affrication is so pervasive that nowadays children have problems at school in writing even common words such as *penso* ‘(I) think’, *salsa* ‘sauce’, *borsa* ‘bag’, *polso* ‘wrist’, *perso* ‘lost’, which are often written as <penzo, salza, borza, perzo, polzo, perzo>.¹² Note the fact that the process may give rise to the neutralization of the phonemic contrast between /s/ and /ts/. For instance, the two words *salsa* ‘sauce’ /'salsa/ and *Salza* /'saltsa/ ‘the name of a Pisan cake shop’ collapse in a unique phonetic form, i.e. ['saltsa], in Tuscan pronunciation. Similarly, the distance between some lexemes may decrease. Consider for instance *orso* ‘bear’, SI /'orso/ and *orzo* ‘barley’, SI /'ɔrdzo/: the two words are pronounced ['ortso] and ['ɔrdzo], respectively, both with an affricate, though with a different voicing value.

The acoustical analysis by Turchi & Gili Fivela (2004) has shown that stop closure is short and less strong in the case of /s/ becoming an affricate with respect to an underlying /ts/. As a matter of fact, the fine phonetic detail (in the sense of Local 2003) of a few milliseconds of difference in the stop closure can easily be recognized at the articulatory and acoustic levels, without having any clear correspondence at the perception level: both underlying /ts/ and the affricate stemming from post-sonorant /s/ are actually perceived as the same sound, with the consequent suspension of a phonemic contrast.

11. The term ‘near-merger’, coined by Labov (1966), has been deeply discussed by Labov himself (1994: 349–370). He underlines the disbelief of the phenomenon by linguists and phoneticians, despite the wide range of data supporting the idea of partial neutralizations or near-mergers. From the theoretical point of view, it is quite obvious that traditional phonology cannot recognize the notion of near-merger, inasmuch as it challenges the principle of binary opposition as well as the symmetry between production and perception.

12. With respect to the affrication process of /s/, a very nice example may be quoted here: in the object of an e-mail message written on November 2010 by a University caretaker of the Department of Linguistics of Pisa University, the following phrase was written: *sospensione della didattica in data 30 Novembre 2010* (i.e. ‘Lessons are suspended on November 30th 2010’), instead of *sospensione*. Who wrote the message was unable to identify the phoneme corresponding to the segment [ts] in a post-consonantal context: is it /s/, to be written as <s>, or /ts/, to be written as <z>? Nowadays, this is the Hamlet’s question for many Tuscan speakers, listeners and writers.

Nowadays, the process of near-merging of /s/ and /ts/ is indeed dynamic. Although we cannot know whether and when a complete merger will succeed, we believe that the preservation of the original sibilant in this context will not succeed.

5.6 *l*-velarization

In the Tuscan varieties of Pisa and Leghorn, a velarization process of the lateral consonant has been observed in context of gemination or in syllable coda position; for instance, *anello* ‘ring’ [aˈnɛːo], *cavallo* ‘horse’ [kaˈvaːo], *molto* ‘much’ [ˈmolto] (Marotta & Nocchi 2001; Nocchi & Marotta 2003).

From a typological point of view, we should observe that velarization of a lateral sound in coda position is not a marked phenomenon. In many languages, there are velarized allophones for /l/, as in Russian or English; also in other Italian dialects the phenomenon is well attested (see Grassi et al. 1997, 2006).

The process of *l*-velarization in Pisa and Leghorn appears to be strongly constrained by sociolinguistic factors, because it is sensitive to age, gender and social class: it occurs more often in young male people belonging to low social classes (see further paragraphs).

In descriptive terms, its *SHAPE* can be summarized as follows: /l/ velarizes when it is geminate or it occupies the coda position in the syllable of a word. In more formal terms, the *SHAPE* of the process is quite simple:

$$(13) \quad /l/ \rightarrow [ɭ] \quad / _ \\ \quad \quad \quad | \\ \quad \quad \quad [\text{Coda}]_o^{13}$$

This phonological process appears to be sensitive to the nature of the preceding segment too: the degree of velarization is inversely proportional to the height of the preceding vowel. Therefore, it is higher when a low vowel precedes, as in the cases of /a/, /ɛ/, /ɔ/, whereas it is lower after a high vowel, especially if it is the front vowel /i/. The acoustic parameter taken into account is the value of F2 (Marotta & Nocchi 2001; Nocchi & Marotta 2003).

We might wonder whether this last segmental constraint must be included in the *SHAPE* of the velarization process. The answer is negative. As a matter of fact, we here are facing a kind of fine-grained phonetic variation: since the lateral becomes progressively more velarized as an inverse function of vowel height,

13. The reference to the syllabic structure is indeed sufficient to capture both contexts of the process, since a geminate consonant is heterosyllabic in Italian (as usual in natural languages), thus associated to two skeletal positions, the first one in coda and the second in onset of the subsequent syllable.

the phenomenon appears to be gradual and continuous, as is often the case with many sociophonetic indexes. Gradual phonetic processes are favored in some specific positions, whereas they are disfavored or weakened in others. In the case of *l*-velarization, typological studies have shown that the prototypical position for [ɭ] to emerge is the coda position; on the other hand, the relation between low vowels and velarization of the lateral is expected because of normal coarticulatory effects on speech production. For this reasons, the syllabic constraint allows to identify the primary context of application of the process, whereas the reference to the preceding segment only explains the gradualness of the process. In other words, the coarticulatory effects apply to the lateral consonant whatever its place of articulation, whereas [ɭ] is produced only in a specific syllabic context.

Therefore, the two constraints holding on *l*-velarization, i.e. vowel height and coda position, do not have the same value in terms of *SHAPE*. Only the syllabic constraint is relevant in a system-oriented perspective, because the syllable position is the context triggering the process, whereas the vowel height constraint has the surface effect of increasing the degree of velarization. The latter constraint exhibits fine phonetic detail which is not relevant for the *SHAPE* of the sociophonetic process (see also the final section of the paper).

5.7 Oxytone infinitives

At the morpho-phonological level, a rather common process taking place in Tuscan varieties of Italian (and in many other sub-standard varieties of the Centre-South of Italy) is the truncation of the verbal infinitive forms, as in the examples in (14):

- (14) Tsc. *cantà* 'to sing', *andà* 'to go', *vedé* 'to see', *sentì* 'to hear', *uscì* 'to go out', *conósce* 'to know', instead of SI *cantàre*, *andàre*, *vedére*, *uscire*, *sentire*, *conoscere*.

The latter example additionally shows that also pro-paroxytones, and not only paroxytones may undergo the truncation process, thus giving rise to paroxytons truncated infinitives. The loss of the final unstressed syllable *-re* of the infinitive is attested across Tuscany as well as in a wide area of the Center-South of Italy (see Rohlf's 1968: 359–360; Savoia 1990). The process of syllable deletion is very frequent in current speech of Tuscan people, especially during informal conversations of the speakers from the North-Western area of the region (Marotta 2000).¹⁴

Its *SHAPE* can be described in the following formal terms:

14. In Standard Italian, especially in the poetic register – but sometimes also in Tuscan rural speech – the verbal infinitives also show allomorphs ending with the rhotic consonant, i.e. with the loss of the final unstressed vowel only, such as in *cantàr*, *vedér*, *sentir*, etc.; see Marotta (2000) for further details.

(15) /-re/ → Ø / _] *Verb Infinitive* #

Oxytone infinitives feed the process of RF (see §5.4). Therefore, instead of sequences such as SI *veder(e) bene* ‘to see well’, *sentir(e) forte* ‘to hear distinctly’, *mangiare tanto* ‘to eat a lot’, in Tuscany we find the corresponding truncated forms: Tsc [ve.de 'b:ɛ:ne] [sen.ti 'f:ɔ:te], [man.dʒa 't:anto], respectively.

The lexicalization of the oxytone forms is demonstrated by their occurrence in the prepausal position (e.g., *ci vuole andare* ‘(s/he) wants to go there’, Tsc [tʃi 'vɔ:le an'da]; *ti ho detto di uscire* Tsc ‘I told you must go out’, Tsc [tɔ d:et:to d u'ʃi]) as well as in hiatus (e.g., *dover(e) andare* ‘to have to go’, Tsc [do,vɛ an'da]; *far(e) entrare* ‘to allow (someone) to in’, Tsc [fa en'tra]). Furthermore, the lack of /s/ affrication after the truncated infinitives confirm that these forms belong to the competence of the Tuscan speakers as such: for instance, *aver sentito* ‘to have heard’, Tsc [a,vɛ s:en'ti:θo], not *[a,vɛ t:sen'ti:θo]; *esser(e) sinceri* ‘to be sincere’, Tsc [ɛs:e sin'tʃe:ri], not *[ɛs:e tsin'tʃe:ri] (Marotta 2000: 196–197).

6. SIZE (or VOLUME)

The *SIZE* (or *VOLUME*)¹⁵ of a sociophonetic variation is the degree of pervasiveness of the process in the phonological system. Basically, a great *SIZE* refers to a large number of segments involved in the process; on the other hand, a small *SIZE* means that only a few segments are involved by the process. In a language variety, the scope and the potential basin of application of a sociophonetic process can therefore be very different.

In particular, if we consider the Tuscan phonological processes discussed in the preceding paragraph with reference to *SHAPE*, we see that their value of *SIZE* can vary in a remarkable manner. *SIZE* appears to be particularly large in the case of RF, because this process applies to every consonant of the phonological inventory by principle. Quite large is also the value of this parameter for *gorgia* as well as for apocope, since in both cases the process entails many segments, that is all the plosives in Florence as well as in the centre of the region (i.e., /p t k b d g/), and all vocalic segments, respectively.

The other phonological processes under examination, i.e. de-affrication, s-affrication and l-velarization exhibit a small *SIZE*, because they are limited to

15. If we look at the properties that physics assigns solids, volume is the proper term to use. However, our metaphor is effective inasmuch as it covers the general way of thinking adopted by the speakers and the vocabulary used by normal people, rather than the scientific register of physics. Therefore, we will consistently use *SIZE* instead of *VOLUME*.

two (in the case of de-affrication) or one segment only. In the case of the oxytone infinitives, the process concerns the morpho-phonological level: only a specific morphological ending triggers the process of syllable deletion.

In conclusion, *SIZE* refers to the scope of the sociophonetic variation with respect to the overall system composed of the phonological categories and their possible variants. We will show that *SIZE* strictly interacts with *THICKNESS* (see §7), since the more the segments involved in a sociophonetic alternation, the greater the relative awareness of the speakers.

SIZE may also refer to the effects of a sociophonetic process on the phonological system. In a system-oriented model such as the one that we are proposing, this second meaning of *SIZE* may become even more important than the first one, inasmuch as the phonetic outputs of a sociophonetic alternation interact with the other elements belonging to the phonological inventory of the language. Therefore, the same sociophonetic process can be assigned different values with regard to the parameter of *SIZE*. For instance, the process of *s*-affrication receives a small *SIZE* according to the first meaning, since it involves one segment only; however, according to the second meaning, it is defined by a very large *SIZE*, because it gives rise to the merging of two different phonemes (i.e. /s/ and /ts/) in a specific context. This merging carries a potential cognitive confusion in the recognition of categories by the speaker-listeners of Central and Southern varieties of Italian.

In conclusion, with regard to this second interpretation, the value assigned to the Tuscan phonological processes considered so far will in general be very low (with the only exception of *s*-affrication), since the surface sociophonetic variation does not conflict with the phonological categories. For instance, the output of *gorgia* produces a set of fricatives which do not merge with any other existing category. *Variatis variandis*, the same is valid for *l*-velarization, apocope and RF.

7. *THICKNESS*

While the two geometrical properties considered so far have absolute geometrical values, i.e. they are independent from the observer, the third property, *THICKNESS*, entails both a physical and a perceptive dimension, inasmuch as it cannot be evaluated without the specific point of view of the observer. *THICKNESS* can be defined as the part of the solid behind its front section for a given subject who is observing it.

In sociolinguistic variation, *THICKNESS* should refer to the degree of control the speaker may have on his/her production with reference to a given sociophonetic process. In other terms, the degree of *THICKNESS* is directly proportional to the percentage of occurrence of a sociophonetic process in everyday speech. In general, the generalization in (16) can be assumed:

(16) if x is more frequent than y , then x is thicker than y

where x and y correspond to two sociophonetic processes characterized by a different percentage of use.

In this sense, the capacity of the speaker of controlling his/her own pronunciation is assumed to be a function of the occurrence of the relevant sociophonetic indexes. Said in a different way: the more the control realized by the speaker, the less frequent the occurrence of sociolinguistically marked variants. Speakers may control their pronunciation in the case of some processes, whereas in others they have no control at all in the production of the phonetic cues indexing a specific sociolinguistic value.

The speaker, his behaviour and his attitudes are crucial for assigning a *THICKNESS* value to a certain sociophonetic process. While *SIZE* and *SHAPE* are attributes of the phonological processes independent from the speaker, *THICKNESS* cannot dispense with the speaker. Consider again the case of *gorgia toscana* and its target classes, i.e. plosives and palatal affricates (§5.1 and §5.2). There seems to be a different behaviour in the speakers with reference to the consonant classes involved. The picture which has been drawn after years of empirical analysis is such that the spirantization of the palatal affricates is less prone to the direct control by the speaker as well as to the censure of the listeners than the spirantization of the stops (Marotta 2008). In our terms, we could say that the first process is thicker than the second. Indeed, native speakers of Tuscan varieties can only pronounce an intervocalic palatal affricate /tʃ/ or /dʒ/ with great articulatory effort and a high degree of self-control.¹⁶ In the same phonetic context, the fricative allophones of the stops can be more easily suppressed, especially at some diaphasic levels.

In the everyday speech of Tuscan speakers almost mandatory is the occurrence of RF as well as of *s*-affrication, with consequent high levels of *THICKNESS*. In the case of RF, the fact that the phenomenon also belongs to Standard Italian, or at least to the Central and Southern varieties of it, encourages the application of the process.

If *gorgia*, RF and *s*-affrication obtain the highest values of *THICKNESS*, the other sociophonetic processes of Tuscan considered here are assigned a middle score, in the case of truncated infinitives, or a relatively low score, in the case of *l*-velarization and apocope.

16. Many Italians probably remember the anecdote that one of the past presidents of the Italian Republic, Carlo Azeglio Ciampi, who was a native speaker of Leghorn, was unable to produce the palatal affricates in intervocalic position (nor could he read it aloud), whereas, in the same context, he could pronounce the intervocalic stops in alternation with their weak counterparts, i.e. the fricative consonants.

In particular, the process of *l*-velarization exhibits a variable degree of *THICKNESS*, since its constraints are twofold: lexical (frequency of the word) and phonetic (vowel quality; see §5.6). With reference to a gradual scale of *THICKNESS*, the picture can be summarized as follows:

- the highest value has to be assigned to words such as *bello* ‘nice’, *fratello* ‘brother’, *sorella* ‘sister’, which are very frequent and contain a geminate lateral and a low [æ] vowel (which is one of the strongest *shibboleth* of Leghornese; Calamai 2004);
- an intermediate value has to be assigned to words such as *polpo* ‘octopus’, *palmo* ‘palm’, *facoltà* ‘faculty’, which are also relatively frequent and contain a low or a mid back vowel;
- the lowest value has to be assigned to words such as *birillo* ‘pin’, *grullo* ‘foolish’, *fanciullo* ‘child’, which are relatively infrequent and contain high vowels.

In conclusion, *THICKNESS* may be considered as a parameter for evaluating the degree of robustness of a sociophonetic index, according to the implication in (17):

- (17) Sociophonetic index:
 more difficult to control → more frequent → more robust → *thicker*

8. *WEIGHT*

Coming back to our metaphor of solids and to the theory of solids in physical geometry, a new analogy may be recognized: in physics, solids have mass and weight, which are properties belonging to them as objects embedded into the physical world. The abstract conception of solids does not entail the property of *WEIGHT*, which becomes indeed necessary where a solid is inserted in a physical environment. In a similar way, sociophonetic indexes do not have *WEIGHT per se*, while they receive a different *WEIGHT* in their usage in everyday speech. The linguistic community assigns the *WEIGHT*. The values of this parameter can be different according to the social evaluation of a given phonetic cue.

In some way, *WEIGHT* appears to be the strongest sociolinguistic property in the proposed set of parameters, because it refers to the social value given to the sociophonetic indexes by the speakers of a linguistic community. What I call *WEIGHT* is basically equivalent to the term ‘prestige’, traditionally employed in sociolinguistics (Trudgill 1972; Labov 2001:60, 196 ff.). As is well known, the sociolinguistic prestige of a variable is directly proportional to its use in more formal styles. With respect to *WEIGHT*, the attributes of heavy and light may easily correspond to covert-prestige and overt-prestige, respectively. In our opinion,

it is beyond dispute that different levels of social salience exist on the perceptual side of the sociolinguistic variation in every speech community. The parameter of *WEIGHT* would precisely represent this perceptual and social salience as expressed by the listeners.¹⁷

In language, in general, and in sociolinguistic environment, in particular, *heavy* can easily be read as synonymous of dialectal, peasant, rustic. In short, it has a negative social connotation. On the opposite side, *light* is positive, as it is synonymous of standard, urban, refined, elegant. Therefore, *heavy* implies [–prestigious] whereas *light* implies [+prestigious].

Coming back to the sociophonetic processes considered so far, *s*-affrication in post-consonantal context is a process which attains a high degree of *THICKNESS* and a low value of *WEIGHT*. The quantitative analysis we carried out on various corpora of spontaneous speech has shown that Tuscan speakers produce [ts] instead of [s] with very high percentages, normally more than (85%).¹⁸ At the same time, speakers who produce *s*-affrication do not appear to be aware of the process. This lack of awareness is the reason for assigning the [+light] value of *WEIGHT*, at least within the geographical boundaries of Tuscany. Outside the region, the values of *WEIGHT* may change. In particular, Tuscan pronunciations such as ['saltsa] 'sauce' or ['bortsa] 'bag' could be evaluated as odd and dialectal, then socially low or even rude (in our terms, as [+heavy]), by Northern speakers of Italian, whereas Southern speakers would probably judge the same pronunciations as normal and similar to the standard, because the same process of *s*-affrication occurs in their varieties of the national language.

17. Passing from one metaphor to another, a quite interesting correlation might be observed: the possible semantic associations concerning the opposition between *heavy* and *light* are all represented in terms of *positive* versus *negative* poles. As a matter of fact, *heavy* has a negative connotation, especially in our contemporary Western society, where *heavy* is associated to fat or low-educated and rude behaviour. On the other side, *light* has a positive connotation: in advertising, in movies as well as in variable performance, people have to be light in their bodies as well as in their behaviour (food must be light, our way of walking and maybe our way of thinking should be light as well, etc.).

18. We are basically referring to the unpublished report by M.A. student Alice Idone in 2010–2011, who carried out quantitative and qualitative analyses on a set of free conversations among young Tuscan people. This research on the occurrence of some phonological processes in Tuscan speech was part of her stage at the Laboratory of Phonetics of the University of Pisa during the academic year 2010–2011. The results reported in Idone's study agree with other empirical data collected and discussed in some M.A. theses of the Master Course in Linguistics at the Department of Linguistics of the same University under the supervision of the author of this paper.

On the opposite site, *l*-velarization occurring in Leghorn and Pisa is a process whose *WEIGHT* receives the value [+heavy] nowadays as well as many years ago. A low prestige value has been assigned to this process from the very beginning of occurrence and it will probably maintain the same *WEIGHT* value in the future. Pronouncing words such as *bello* ‘nice’ or *cancello* ‘gate’ with [ɫ] is a clear sociolinguistic marker, inasmuch as it allows the Tuscan listener to assign a low social and educational status to the speaker. This means that *l*-velarization is a sociophonetic index which performs a steep slope in style shifting, as it normally happens in the case of heavy values of *WEIGHT*. The same happens in the sociolinguistic markers proposed by Labov (2001: 196 ff.), which may even become stereotypes.¹⁹ Despite its sociophonetic *WEIGHT*, the velarized lateral appears to be currently spreading across North-Western Tuscany, as it occurs not only in the cities of Pisa and Leghorn, but also in the areas of the countryside close to Pisa and Leghorn, as well as in the surrounding Tyrrhenian coast. In the modern social stratification of contemporary Tuscany, overt norms, reflecting the standard manner of speaking, can be balanced by covert norms, which assign a positive value to the non-standard forms used by people in everyday life. Indeed, one of the principles set out by Labov (2001: 196) perfectly applies in the case of *l*-velarization in Tuscan varieties: “every overtly stigmatised feature has prestige in the social contexts where it is normally used”. Therefore, some social groups, such as the people working in the docks of Leghorn, use *l*-velarization as a cue for marking their social identity and as a tool for expressing the sense of belonging to a special community and a well defined social network, where strong bindings and close-knit networks are holding (L. Milroy 1987; J. Milroy 1992). On the other hand, such a heavy index (or marker) can strongly be stigmatised by the educated and upper class speakers of the cities of Leghorn and Pisa, on the behalf that “every prestige feature will be awarded an equal and opposite stigma in those opposing contexts” (Labov 2001: 196).

At present, there seems to be a change in progress in the *WEIGHT* values of *gorgia Toscana* in Italy. This process normally occurs in the everyday speech of Tuscan speakers, without any constraint related to the social status of the speaker or his/her education, because it is a thick and very robust feature of ‘Tuscanicity’. As we saw, its *SIZE* is great and its *THICKNESS* is high. Until some years ago, *gorgia toscana* could be assigned the *WEIGHT* value [+light], then [+prestigious]. Nowadays, there are some signs that its *WEIGHT* is becoming heavier than before, at least outside

19. The values of *WEIGHT* for velar /ɫ/ in North-Western Tuscan varieties are not directly comparable with those of dark /ɫ/ occurring in some varieties of British English, because different opinions are reported in the literature about the prestige of this allophone. Some references on the matter are Wells (1982), Horvath & Horvath (1997), Tollfree (1999).

Tuscany. This change in the evaluation of the process is due to the interaction between internal and external factors in language behavior. Indeed, the Northern varieties of Italian are nowadays perceived as more prestigious than the Tuscan ones,²⁰ and this has two consequences at least.

On one hand, the phonetic features of Northern varieties of Italian have recently acquired an increasingly prestigious sociolinguistic value (for instance, *s*-voicing in intervocalic position); on the other hand, some sociophonetic indexes typical of Tuscany are slowly losing their traditional sociolinguistic prestige, at least outside the region. Therefore, it is worthwhile to underline that the *WEIGHT* of *gorgia* is changing from [light] to [heavy] outside Tuscany, whereas no change in the evaluation of the process seems to occur within the regional boundaries. In particular, in the case of Tuscan speakers, especially Florentine, sociolinguistic variables such as education and social class do not play any role. As a matter of fact, recent studies carried out with the matched-guise technique have shown that the variety spoken in Florence is still perceived as standard and more prestigious by the speakers of other Tuscan areas (Calamai & Ricci 2005; Calamai 2011; Biliotti & Calamai 2012). In particular, the variety of Florence is traditionally assigned overt prestige, as associated to positive attributes like elegance, culture and tidiness. On the other hand, in Tuscany, multiple competing norms are also active: alongside the Florentine norm, which applies more or less to the entire region, specific and different local norms are perceived as prestigious in almost every capital of the regional districts (Cravens & Giannelli 1995; Pacini 1998; Pacini & Giannelli 1999).

The awareness of an increasing *WEIGHT* value, with a consequent loss of prestige outside Tuscany, may induce the speakers, especially the youngest ones, to try to control their production, then reducing the degree of *THICKNESS* of the phenomenon. Two forces are at play here: *WEIGHT* and *THICKNESS*. The result can be a reduction of both parameters, since speakers could try to produce a phonetic target they cannot reach, i.e. a plosive consonant in intervocalic context, by producing a segment closer to the target, such as a semifricative²¹ or a fricative. For instance, in Pisan or Leghornese speech, we have found speakers who produce [x] for /k/, then counterbalancing the typical trend towards the deletion of the velar

20. This picture was already drawn in the sociolinguistic analysis carried out by Galli De' Paratesi (1984) roughly thirty years ago. She showed a clear trend towards the spreading of Northern features in the speech of Florentine young speakers, as, for instance, *s*-voicing in intervocalic context. For further studies on the matter, not restricted to the Tuscan area, see Baroni (1983), Volkart-Rey (1990), Bernhard (1998).

21. By 'semifricative' we mean a segment with a stop closure followed by a long VOT without any sign of spike on the spectrogram; see Marotta (2008) for the phonetic details.

plosive when preceded by a vowel (Marotta 2001/2004). Therefore, a production such as [la 'xa:sa] 'the house' is perceived as more prestigious, or in our terms, lighter than [la 'a:sa], with /k/ deletion.

In a parallel way, in North-Western Tuscany, saying [kwəl'kɔ:za] 'something', [na:zo] 'nose', [pi:za] 'Pisa' may be considered nowadays more prestigious than [kwəl'kɔ:sa], [na:so], [pi:sa] (with voiceless [s]), especially by the youngest speakers, because [z] is the phonetic output of the sibilant in intervocalic context in the Northern pronunciations of Italian. The use of allophones considered more prestigious than the local ones by the speakers who want to attain a better social and education level is widely described in traditional sociolinguistic analyses, and Tuscany makes no exceptions.²²

9. Interactions among parameters

In this section, the possible interactions among the proposed parameters will be shortly presented.

First, there seems to be a conspiracy between *SIZE* and *THICKNESS*: the more the segments involved in a sociophonetic process, the higher the frequency of usage of the indexes and the higher the degree of *THICKNESS*. On the other hand, *SIZE* does not directly interact with *WEIGHT*. We have seen that processes with great *SIZE* can have a light value in social evaluation (e.g. *gorgia*), whereas others, despite their small *SIZE*, can be judged as heavy (e.g. *l*-velarization).

Although *THICKNESS* interacts with *WEIGHT*, there is no direct and proportional relation like the one summarized in the formula in (18):

(18) if a process is thicker, then it is heavier.

In fact, among the sociophonetic processes occurring in Tuscany, some have a high degree of *THICKNESS* (e.g. *gorgia*, in both its facets of stop lenition and deaf-frication of palatal consonants; see also *s*-affrication) and they still maintain a low value of sociolinguistic *WEIGHT*. Some others (e.g. truncated infinitives) show an increasing degree of *THICKNESS* which is strictly related to greater *WEIGHT*.

A further theoretical aspect concerns the nature of the sociophonetic parameters proposed: they do not exhibit the same nature. *SHAPE* and *SIZE* are

22. These phenomena have traditionally been considered as *ipercorettismi* in Italian traditional dialectology. The meaning of the term *ipercorettismo* seems to be very similar to that of *ambition* as it has been recently proposed in sociophonetic theory, especially in the German framework; see for instance Nocchi & Filipponio (2011).

descriptive parameters; they belong to the linguistic system more than to the speaker-listener of a speech community. On the other hand, *THICKNESS* refers to the speaker's behavior, whereas *WEIGHT* makes crucial reference to the listener and to social evaluation.

Focusing on the perceptive corner of sociolinguistic variation, listeners may give different evaluations of the same sociophonetic index. Differences are ultimately based on diatopic variation: in order to assign a value for the parameter of *WEIGHT*, the origin of the listener is crucial. With reference to the Tuscan sociophonetic processes we have discussed so far, there are at least two kinds of listeners, i.e. Tuscan listeners and non-Tuscan listeners, with subsequent potential differences in the evaluation of the same variable. Therefore, the perception as well as the sociolinguistic evaluation of phonetic indexes may be different and even opposite in the same country. For instance, *gorgia toscana* has a light value of *WEIGHT* not only for Tuscan listeners, but also for people speaking a Central or Southern variety of Italian. In parallel, the process is becoming heavier for Northern listeners, who could even consider the *gorgia* as a Tuscan stereotype, then negatively judged. A partly similar picture can be obtained for *s*-affrication: many varieties of Centre-South of Italy share this process with Tuscany, whereas Northern varieties lack it. Therefore, the perception and the consequent social evaluation may be different: [+light] in the first case, [+heavy] in the second one.

A parameter seems still to be necessary in the model, which can describe the effects of sociophonetic indexes on the phonological system. Some aspects of this topic are captured by *SIZE*, inasmuch as it makes reference to the number of segments enrolled by the process. However, as we already have underlined (see §6), the impact of sociophonetic variation on the phonological structure has to be considered too. In this respect, *s*-affrication is particularly revealing. *S*-affrication has a very small *SIZE* facing with a strong structural relevance, because of its impact on the phonological competence of the speakers. Among the processes we have considered, *s*-affrication is actually the only one showing a remarkable phonological relevance, thus enlightening possible scenarios for linguistic change.

As a final remark, we would like to observe the rather stable production of the vocalic segments with respect to the high variability shown by the consonants in Tuscan varieties. Of the seven phonological processes considered, only one strictly involves vowels (i.e. apocope), another refers to a syllable unit (oxytone infinitives), whereas the remaining five have a consonant as target. All processes but RF are weakening processes. Therefore, the phenomena investigated can ultimately be interpreted in the light of a general articulatory strategy, a special 'manner of speaking' typical of Tuscan speakers. This special speech quality appears to be crucially marked by the feature of laxing (Marotta 2001/2004).

10. Discussion

The parameters we have proposed exhibit a qualitative and discrete nature, instead of a quantitative and gradual one, as is more usual within the framework of sociophonetics.

Some general remarks are then needed in order to show the possible advantages of adopting such model of sociophonetic variation. A first benefit derives from the lack of redundancy: each parameter is distinctive and independent from the others. In other terms, it is not possible to predict the value for a parameter *x* from the one relative to a parameter *y*. At the same time, there is not a strict correlation among the different parameters, but only a lax relation (for instance, between *THICKNESS* and *WEIGHT*).

Furthermore, the metaphorical parameters allow an easier and more transparent comparison between the various sociophonetic indexes: a process *a* may receive a higher value than a process *b*, but a lower value than a process *c*, with respect to a specific parameter. In such a way, the adoption of a point of view based on the notion of discreteness permits to directly compare the sociophonetic indexes, with the result of being able to forecast their spreading or compression in space and time.

Despite their qualitative and discrete nature, our parameters could be represented in terms of a multi-factorial scale. In particular, with reference to the typical phonological processes occurring in the Tuscan varieties considered so far, each parameter could be assigned a different value going from a minimum up to a maximum. However, in order to decide the exact numerical value to assign to each parameter with respect to the phenomena considered, we would need new and more accurate analyses, especially on the perceptual side. Therefore, at least for the moment, we prefer to dispense with explicit multi-factorial scales.

Our model tries to conjugate the fine-grained description of phonetic events with the adoption of a systemic analysis, thus leaving aside any theoretical perspective exclusively oriented towards the surface outputs of the single speaker's behavior. An example of such surface-oriented perspective is the one based on Exemplar Theory which is often enrolled by socio-phoneticians (e.g., Johnson 1997; Foulkes & Docherty 2006; Carlson & Hawkins 2007). As a matter of fact, the basic tenet of Exemplar Theory is that phonetic categories map directly onto phonological categories; the basic elements of our phonetic knowledge are relational, dynamic, self-organizing and entirely context-sensitive (Pierrehumbert 2001, 2006; Hawkins 2010); consequently, it seems that we do not need phonology any more. In this approach, the cognitive process of evaluation and classification is driven by the degree of resemblance between the single and concrete manifestations of the category in a word and the members assumed as prototypical for a

given category. Therefore, the same categories are viewed as a set of single memorized repetitions, the so-called exemplars. Only individual properties, which are neither abstract nor rule-feeding, are stored and represented in the mind. Human subjects memorize all the linguistic information (lexical, morphological, phonetic) in an additive manner, via single *items*, i.e. *verbatim*.

Following Labov (2006), we argue that such an approach cannot really account for the relatively abstract and symbolic categories holding in language structure. In our opinion, Labov (2006) is right in maintaining a discrimination among phonetic (i.e. physical, concrete) elements and phonemic (i.e. formal, abstract) units. However, this leads us to a thorny theoretical problem that cannot be exhaustively debated here, so we leave the question open for future discussion.

11. Conclusion

A quite surprising finding of the research presented in the *Atlas of North American English* by Labov et al. (2006) was that the regional varieties of English in North America continue to diverge. The common and really naïve assumption that dialects should disappear in our contemporary age, due to shared education, mass media communication, high mobility of people and so on has then been falsified by the empirical data. Speakers of the third millennium still encode social-indexical information in their speech. With such a behavior, they are able to project their own identity both inside and outside the speech community they belong to (see Tabouret-Keller 1997).

Si parva licet, we might say that the same pattern emerges nowadays in Tuscany: Tuscan speakers do still show their peculiar sociophonetic indexes, and sometimes they are proud of them, especially of some, like *gorgia toscana*. They do not want to lose their cultural identity as well as their way of speaking. Sometimes they are aware of the phonetic cues which give them the status of Tuscan people, sometimes they are not. In any case, the sociophonetic indexes are alive, and nothing seems to indicate that their life should be a short one. Social and psychological factors such as identity and attitudes are therefore confirmed to be strong forces holding in the speaker's phonetic performance.

A comparison of our parameters with the three classes of socially-marked variables proposed by Labov (2001: 196–197), i.e. indicators, markers and stereotypes, has finally to be handled with. As is well-known, the Labovian classes form a sort of chain, having different and increasing degrees of salience and awareness with respect to the members of a speech community. The indicators are linguistic variables distributed among the social groups of a community which use them without any reference to change in style; these variables are normally employed

with zero degree of social awareness by the speakers, since they are “never commented on or even recognized by the listeners”. By contrast, the markers are linguistic variables which have acquired a social recognition (usually in the form of social stigma) and which show a consistent and layered structure along the diastatic and the stylistic dimensions; speakers select one variable or another, according to the formality of the communicative context. Finally, the *stereotypes* are linguistic variables strongly marked and normally “overt topics of social comment”; they are employed by some special groups of the speech community, usually of a low socio-economic level; speakers producing stereotypes cannot be aware of realizing a stereotype form themselves, whereas speakers of a higher social level clearly exhibit a stigma towards them.

With respect to the parameters taken from the metaphor of solids, degree of awareness and context sensitivity may be covered by the parameter of *THICKNESS*, whereas the dimension of *WEIGHT* is mostly represented by the behavior of the stereotypes, although it might be on the background of the other two classes too.

Some doubts have been casted on the discreteness of these Labovian classes as well as on their corresponding levels of salience (see for instance Docherty 2007: 22). In our opinion, discreteness is needed, in sociophonetic analysis too, if we do not want to run the risk of losing ourselves in the wide sea of surface variation. The metaphorical parameters presented here should indeed help us in finding out the way to discriminate between the relevant social-indexical information and the irrelevant one. The conceptual nature of the parameters, despite their metaphorical source, is that of descriptive and discrete entities. To express the same concept in the metaphorical terms of *Accademia della Crusca* (a historical institution originally devoted to the preservation of the purity of Italian language), we could say that we propose to do sociolinguistic research *separando il grano dalla crusca*, i.e. “by separating the wheat from the grain”.

References

- Agostiniani, Luciano. 1989. “Fenomenologia dell’elisione nel parlato in Toscana”. *Rivista Italiana di Dialettologia* 13. 7–46.
- Agostiniani, Luciano. 1992. “Su alcuni aspetti del ‘rafforzamento sintattico’ in Toscana e sulla loro importanza per la qualificazione del fenomeno in generale”. *Quaderni del Dipartimento di Linguistica dell’Università di Firenze* 3. 1–28.
- Agostiniani, Luciano & Luciano Giannelli. 1990. “Considerazioni per un’analisi del parlato toscano”. *L’italiano regionale. Atti del XVIII Congresso Internazionale della Società di Linguistica Italiana* ed. by Manlio A. Cortelazzo & Alberto M. Mioni, 219–237. Roma: Bulzoni.
- Albert, Abraham A. 1949. *Solid Analytic Geometry*. Chicago: McGraw Hill Book Company Inc.

- Baroni, Maria Rosa. 1983. *Il linguaggio trasparente. Indagine psicolinguistica su chi parla e chi ascolta*. Bologna: Il Mulino.
- Bernhard, Gerlad. 1998. *Das Romanesco des ausgehenden 20. Jahrhunderts. Variationslinguistische Untersuchungen*. Tübingen: Niemeyer.
- Bertinetto, Pier Marco & Michele Loporcaro. 2005. "The sound pattern of Standard Italian, as compared with the varieties spoken in Florence, Milan and Rome". *Journal of the International Phonetic Association* 35. 131–151. DOI: 10.1017/S0025100305002148
- Biliotti, Francesca & Silvia Calamai. 2012. "Linguistic opinions and attitudes in Tuscany". *Proceedings of "Sociophonetics, at the crossroads of speech variation, processing and communication"* ed. by Silvia Calamai, Chiara Celata & Luca Ciucci, 1–4. Pisa: Edizioni della Normale.
- Borrelli, Doris. 2002. *Raddoppiamento Sintattico in Italian: A Synchronic and Diachronic Cross-Dialectal Study*. New York: Routledge.
- Calamai, Silvia. 2004. *Il vocalismo tonico pisano e livornese. Aspetti storici, percettivi, acustici*. Alessandria: Edizioni dell'Orso.
- Calamai, Silvia. 2011. "Per una storia della pronuncia degli italiani: opinioni e atteggiamenti intorno alla pronuncia fiorentina". *Storia della lingua italiana e storia dell'Italia Unita. L'italiano e lo stato nazionale. Atti del IX Convegno ASLI* ed. by Annalisa Nesi, Silvia Morgana & Nicoletta Maraschio, 15–184. Firenze: Cesati.
- Calamai, Silvia & Irene Ricci. 2005. "Un esperimento di *matched-guise* in Toscana". *Studi Linguistici e Filologici on Line* 3(1). 63–105.
- Carlson, Rolf & Sarah Hawkins. 2007. "When is fine phonetic detail a detail?". *Proceedings of the 16th International Congress of Phonetic Sciences* ed. by Jürgen Trouvain & William J. Barry, 211–214 (ID 1721). <http://www.icphs2007.de/>
- Celata, Chiara. 2008. "I contrasti allofonici nella percezione nativa e non-nativa". *Processi fonetici e categorie fonologiche nell'acquisizione dell'italiano* ed. by Lidia Costamagna & Giovanna Marotta, 147–163. Pisa: Pacini.
- Celata, Chiara. 2009. "The impact of allophonic variation on L2 speech perception". *Recent Research in Second Language Phonetics/Phonology. Perception and Production* ed. by Michael A. Watkins, Andreia S. Rauber & Barbara O. Baptista, 64–79. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Coulmas, Florian. 1998. "Introduction". *Handbook of Sociolinguistics* ed. by Florian Coulmas, 1–11. Oxford & Malden: Blackwell. DOI: 10.1111/b.9780631211938.1998.x
- Coulmas, Florian. 2001. "Sociolinguistics". *Handbook of Linguistics* ed. by Mark Aronoff & Janie Rees-Miller, 563–581. Oxford & Malden: Blackwell.
- Cravens, Thomas & Luciano Giannelli. 1995. "Relative salience of gender and class in a situation of multiple competing norms". *Language Variation and Change* 7. 261–285. DOI: 10.1017/S0954394500001010
- Deshaiés-Lafontaine, Denise. 1974. *A sociophonetic study of a Québec French community: Trois-Rivières*, PhD Thesis, London, University College of London.
- Docherty, Gerard J. 2007. "Speech in its natural habitat: Accounting for social factors in phonetic variability". *Papers in Laboratory Phonology* 9 ed. by Jennifer Cole & José Ignacio Hualde, 1–35. Berlin: Mouton de Gruyter.
- Fanciullo, Franco. 1997. *Raddoppiamento sintattico e ricostruzione linguistica nel Sud italiano*. Pisa: ETS.
- Foulkes, Paul & Gerard J. Docherty. 2006. "The social life of phonetics and phonology". *Journal of Phonetics* 34. 409–438. DOI: 10.1016/j.wocn.2005.08.002

- Galli De' Paratesi, Nora. 1984. *Lingua toscana in bocca ambrosiana. Tendenze verso l'italiano standard: un'inchiesta sociolinguistica*. Bologna: il Mulino.
- Giannelli, Luciano. 1976. *Toscana*. Pisa: Pacini (2nd edition: 2000).
- Giannelli, Luciano. 1988. "Toscana". *Lexikon der Romanistischen Linguistik* ed. by Gunter Holtus, Michael Metzeltin & Christian Schmitt, 594–606. Tübingen: Niemeyer.
- Giannelli, Luciano & Leonardo M. Savoia. 1978. "L'indebolimento consonantico in Toscana, I". *Rivista Italiana di Dialettologia* 2. 25–58.
- Giannelli, Luciano & Leonardo M. Savoia. 1979–80. "L'indebolimento consonantico in Toscana, II". *Rivista Italiana di Dialettologia* 3–4. 39–101.
- Giannelli, Luciano & Leonardo M. Savoia. 1991. "Restrizioni sull'esito [h] da *t* in fiorentino e nelle altre varietà toscane". *Studi Italiani di Linguistica Teorica e Applicata* 20(1). 3–57.
- Giannelli, Luciano & Beatrice Pacini. 1998. "Il processo di cambiamento dell'indebolimento consonantico a Cortona: studio sociolinguistico". *Rivista Italiana di Dialettologia* 22. 15–57.
- Grassi, Corrado, Alberto Sobrero & Tullio Telmon. 1997. *Fondamenti di dialettologia*. Roma & Bari: Laterza.
- Grassi, Corrado, Alberto Sobrero & Tullio Telmon. 2006. *Introduzione alla dialettologia italiana*. Roma & Bari: Laterza.
- Hawkins, Sarah. 2003. "Rules and representations of systematic fine phonetic detail in speech understanding". *Journal of Phonetics* 31. 373–405. DOI: 10.1016/j.wocn.2003.09.006
- Hawkins, Sarah. 2010. "Phonetic variation as communicative system: Perception of the particular and the abstract". *Papers in Laboratory Phonology 10* ed. by Cécile Fougeron, Barbara Kühnert, Mariapaola d'Imperio & Nathalie Vallée, 479–510. Berlin: Mouton de Gruyter.
- Holtus, Gunter, Michael Metzeltin & Christian Schmitt, eds. 1988. *Lexikon der Romanistischen Linguistik, IV: Italienisch, Korsisch, Sardisch*. Tübingen: Niemeyer.
DOI: 10.1515/9783110966107
- Horvath, Barbara M. & Ronald J. Horvath. 1997. "A multilocality study of a sound change in progress: The case of *l* vocalization in New Zealand and Australian English". *Language variation and Change* 13(1). 37–57. DOI: 10.1017/S0954394501131029
- Jannedy, Stephanie & Jennifer Hay. 2006. "Modelling sociophonetic variation". *Journal of Phonetics* 34. 405–408. DOI: 10.1016/j.wocn.2006.08.001
- Johnson, Keith. 1997. "Speech perception without speaker normalization: An exemplar model". *Talker variability in Speech Processing* ed. by Keith Johnson & John W. Mullennix, 145–165. San Diego: Academic Press.
- Labov, William. 1966. *The Social Stratification of English in New York City*. Washington DC: Center for Applied Linguistics.
- Labov, William. 1972. *Sociolinguistic patterns*. Oxford: Blackwell.
- Labov, William. 1975. "On the use of the present to explain the past". *Proceedings of the 11th International Congress of Linguists, Vol. I* ed. by Luigi Heilman, 825–251. Bologna: il Mulino.
- Labov, William. 1994. *Principles of Linguistic Change, vol. I, Internal Factors*. Oxford: Blackwell.
- Labov, William. 2001. *Principles of Linguistic Change, vol. II, Social Factors*. Oxford: Blackwell.
- Labov, William. 2006. "A sociolinguistic perspective on sociophonetic research". *Journal of Phonetics* 34. 500–515. DOI: 10.1016/j.wocn.2006.05.002
- Labov, William, Sharon Ash & Charles Boberg. 2006. *The Atlas of North American English*. Berlin: Mouton De Gruyter.
- Local, John. 2003. "Variable domains and variable relevance: interpreting phonetic exponents". *Journal of Phonetics* 31. 321–339. DOI: 10.1016/S0095-4470(03)00045-7

- Loporcaro, Michele. 1997. *L'origine del raddoppiamento fonosintattico. Saggio di fonologia diacronica romanza*. Basel & Tübingen: Francke.
- Marotta, Giovanna. 1983–1986. “Rhythmical constraints on syntactic doubling”. *Journal of Italian Linguistics* 8. 35–52.
- Marotta, Giovanna. 1995. “Apocope nel parlato di Toscana”. *Studi Italiani di Linguistica Teorica e Applicata* 24(2). 297–322.
- Marotta, Giovanna. 2000. “Oxytone infinitives in the dialect of Pisa”. *Phonological theory and the dialects of Italy* ed. by Lori Repetti, 191–210. Amsterdam & Philadelphia: Benjamins.
- Marotta, Giovanna. 2001/2004. “Non solo spiranti. La ‘Gorgia toscana’ nel parlato di Pisa”. *L’Italia Dialettale* 62. 27–60.
- Marotta, Giovanna. 2008. “Lenition in Tuscan Italian (Gorgia Toscana)”. *Lenition and Fortition* ed. by Joaquim Brandao de Carvalho, Tobias Scheer & Philippe Ségéral, 235–271. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110211443.2.235
- Marotta, Giovanna & Nadia Nocchi. 2001. “La liquida laterale nel livornese”. *Rivista Italiana di Dialettologia* 25. 285–326.
- Milroy, Lesley. 1987. *Language and Social Networks* (2nd ed.). Oxford: Blackwell.
- Milroy, James. 1992. *Linguistic Variation and Change*. Oxford: Blackwell.
- Nocchi, Nadia & Giovanna Marotta. 2003. “A phonetic ‘conspiracy’: Low vowels and velarized lateral in Leghorn”. *Proceedings of the 15th International Congress of Phonetic Sciences* ed. Maria Josep Solé, Danile Recasens & Joaquim Romero, 909–912. Barcelona: Casual Productions.
- Nocchi Nadia & Lorenzo Filippino. 2011. “Lo vuoi co[z]i o co[s]i? A sociophonetic study on sibilants in the regional Italian of Livorno (Tuscany)”. *Proceedings of “Sociophonetics at the crossroads of speech variation, processing and communication”* ed. by Silvia Calamai, Chiara Celata & Luca Ciucci, 53–56. Pisa: Edizioni della Normale.
- Pacini, Beatrice. 1998. “Il processo di cambiamento dell’indebolimento consonantico a Cortona: studio sociolinguistico”. *Rivista Italiana di Dialettologia* 22. 15–57.
- Pacini, Beatrice & Luciano Giannelli. 1999. “Norma multipla e dinamiche sociolinguistiche nel consonantismo della Toscana orientale”. *Dialetti oggi* ed. by Giovanna Marcato, 141–153. Padova: Unipress.
- Peirce, Charles S. 1972. *The Essential Writings* ed. by Edward C. Moore. New York: Harper & Row.
- Peirce, Charles S. 1991. *Writings on Semiotics* ed. by James Hoopes. Chapel Hill & London: The University of North Carolina Press.
- Pierrehumbert, Janet. 2001. “Exemplar dynamics: Word frequency, lenition, and contrast”. *Frequency and the Emergence of Linguistic Structure* ed. by Joan L. Bybee & Paul Hopper, 137–157. Amsterdam & Philadelphia: Benjamins.
- Pierrehumbert, Janet. 2006. “The next Toolkit”. *Journal of Phonetics* 34. 515–530. DOI: 10.1016/j.wocn.2006.06.003
- Rohlf, Gerhard. 1968. *Grammatica storica della lingua italiana e dei suoi dialetti, Vol. II, Morfologia*. Torino: Einaudi.
- Savoia, Leonardo M. 1990. “Proprietà vocaliche e struttura metrico-sillabica in alcune varietà abruzzesi”. *Studi in memoria di Ernesto Giammarco* ed. by Luisa Mucciante, 331–380. Pisa: Giardini.
- Silverstein, Michael. 2003. “Indexical order and the dialects of social life”. *Language and Communication* 23. 267–274. DOI: 10.1016/S0271-5309(03)00013-2

- Tabouret-Keller, Andree. 1997. "Language and Identity". *The Handbook of Sociolinguistics* ed. by Florian Coulmas, 315–326. Oxford: Blackwell.
- Telmon, Tullio. 1996. "Varietà regionali". *Introduzione all'italiano contemporaneo, vol. II, La variazione e gli usi* ed. by Alberto A. Sobrero, 93–149. Bari & Roma: Laterza.
- Thomason, Sarah G. & Terrence Kaufman. 1992. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
- Tollfree, Laura. 1999. "South-East London English: Discrete versus Continuous Modelling of Consonantal Reduction". *Urban Voices: Accent Studies in the British Isles* ed. by Paul Foulkes & Gerard J. Docherty, 163–184. London: Arnold.
- Turchi, Laura & Barbara Gili Fivela. 2004. "L'affricazione di /s/ postconsonantico nella varietà pisana di italiano". *Il parlato italiano* ed. by Federico Albano Leoni & Renata Savy, 1–25. Napoli: D'Auria Editore.
- Trudgill, Peter. 1972. "Sex, covert prestige and linguistic change in Urban British English". *Language in Society* 1. 179–195.
- Volkart-Rey, Reinhardt. 1990. *Atteggiamenti linguistici e stratificazione sociale. La percezione dello status sociale attraverso la pronuncia. Indagine empirica a Catania e a Roma*. Roma: Bonacci.
- Wells, John C. 1982. *Accents of English*, Cambridge & New York: Cambridge University Press.

Sound archives and linguistic variation

The case of the Phlegraeen diphthongs

Rosanna Sornicola* and Silvia Calamai**

* Università degli Studi di Napoli “Federico II”

** Università degli Studi di Siena

Sound archives are important resources for sociophonetic analysis: first, they contain relatively uncontrolled speech styles, not usually included in the speech databases used in sociophonetic research; second, they allow us to study in a historical perspective some phonetic phenomena that would otherwise be known only for their most recent or contemporary manifestations. Several complex phonetic phenomena such as Romance diphthongization may be better understood by means of sound archives of spontaneous speech. The paper describes the general principles underlying the building of ADICA (*Archivio dei dialetti campani*), an archive of spoken dialectal texts from the Phlegraeen area. The main features of Phlegraeen diphthongs are thus discussed with particular attention to their variability, their social distribution, together with their historical development.

1. Introduction

The paper aims at illustrating the usefulness of sound archives of spontaneous speech for a better understanding of complex phonetic phenomena which may be of interest in the domain of sociophonetics. It also deals with some problems in the exploitation of sound archives, especially those not originally conceived for sociophonetic analysis.

The paper is organized as follows. §2 illustrates the principles that inspire the construction of a sound archive for spontaneous speech and individual biographies, within a typically European research stream as opposed to the correlational analysis characterizing a large part of Anglo-American sociolinguistics. §3 describes the characteristics of Phlegrean diphthongs and of related phenomena in Romance diphthongization. Since the Phlegrean territory is very complex also

from the historical point of view, some diachronic aspects will also be discussed. §4 deals with the variability of diphthongs as attested in the sound archive under examination. In the final part of this section, two fundamental aspects of a sociophonetic research based on archive data are discussed. First, archives contain relatively uncontrolled, sometimes relaxed speech styles, that are not usually included in the speech databases used in sociophonetic research. Second, archives allow us to study in a historical perspective (specifically, according to the real time paradigm) some phonetic phenomena that would otherwise be known only for their most recent or contemporary manifestations.

2. Sound archives for sociophonetic analysis

Sound archives are important resources for sociophonetic analysis. Oral history interview recordings, ethnographic field and traditional music recordings, vernacular speech, local and regional languages recordings offer an enormous amount of material that can be exploited for the observation of the fine phonetic detail of phonological processes. A renewed interest in sound archives is currently visible in various parts of the world, at least since the UNESCO Convention for the Safeguarding of the Intangible Cultural Heritage of October 17th, 2003. To prevent the irreversible deterioration of analogical storage devices, sound archives are digitalized, catalogued, and published online. Several notable examples can be cited: in Europe, we should mention the activity of the Phonothèque of the Maison Méditerranéenne des Sciences de l'Homme in Aix-en-Provence (<http://phonotheque.mmsh.univ-aix.fr/>) as well as the activity of the Phonogram Archive of the University of Zurich (<http://www.phonogrammarchiv.uzh.ch/>). As for Italy, the project *Grammo-foni, Le soffitte della Voce* (<http://grafo.sns.it>) is aimed at recovering Tuscan sound archives (Calamai 2011; Calamai & Bertinetto 2012). For the most part, the sound archives digitalized in *Grammo-foni* were originally conceived for a variety of purposes, including historical, sociological, anthropological and demological. The majority of the sound archives available for consultation are not originally conceived by sociophoneticians and one may ask whether they can usefully be exploited for the purposes of sociophonetic research as well and, if they can, whether they can benefit sociophonetic research and how. The presence of speech styles normally absent or underrepresented in traditional phonetic analysis is a benefit in itself, especially for those phenomena that are typically present in less controlled speech, as the following sections of this paper will show.

The sound archive collected at the University of Naples since the beginning of 2000 has an intermediate character, with respect to the above, in the sense that it

has been collected by linguists for purposes which are not only for sociophonetics. ADICA (*Archivio dei dialetti campani*) is an archive of spoken dialectal texts from the Phlegrean area (Procida, Monte di Procida, Ischia, Pozzuoli). It contains all the recordings collected by scholars and students working on the sociolinguistics of Campania dialects (Sornicola 2002; Di Salvo 2006).¹ It contains 130 hours of recordings from 120 speakers. The map in Figure 1 illustrates the area in which the recordings have been collected.



Figure 1. The area under investigation.

1. The ADICA project has received funding from Regione Campania. The phase of collecting, digitizing and archiving the data has recently been concluded and the obtained archive will soon be published online. A preview of the archive can be found at the following web site: <http://www.innova.campania.it/newsletter/num11/n1.htm>

Starting from an archive of dialectal speech is of crucial importance for the study of urban dialectology in the southern territories of Italian peninsula. The traditional models of urban dialectology, mostly elaborated in an Anglo-American and German framework, focus on the analysis of variation in the sub-standard language. However, north-American, British and German towns are substantially different from the big urban areas of Southern Italy. The former are interested in processes of linguistic standardization that involve large parts of the social classes. The latter still present high levels of dialectal uses that (i) are employed in functionally different contexts; (ii) are particularly vital in code-switching; (iii) interfere with regional and sub-standard Italian.

Dialectological and ethnographic sound archives may have different goals from those of current sociophonetic research (but it is useful to recall that the corpora of data collected by the French and Swiss dialectologists of the end of the 19th and the beginning of the 20th century prompted the pioneering analysis of linguistic variability that later on was to become a source of inspiration for sociophonetic research). If investigated in the light of the basic principles of variation and variability of European linguistics, they offer a considerable amount of new data for sociophonetics. In particular, ADICA has been constructed according to some general principles that clearly illustrate the potential of combining dialectology and sociolinguistics:

- a. the principle of relativity of variation;
- b. the principle of the centrality of the individual;
- c. the principle of microscopy in the study of variation;
- d. the principle of oscillation and the principle of linguistic context;
- e. the principle of having recourse to spontaneous speech.

The principle of relativity of variation assumes that the range of variation is not defined by one absolute theoretical unit with respect to which all other variants are to be considered as 'alterations'; on the contrary, it is defined by the presence of more than one variant. For this reason, the archiving procedure of all items is based on the underlying etymological lexeme. In this respect, the historical perspective proves the most flexible for storing the multiple outputs of a segment.

The principle of the centrality of the individual speaker assumes that the individual has to be considered as the basic unit of analysis, according to a series of arguments developed by several scholars from Schuchardt to Jespersen, Mathesius and the founders of Romance dialectology. The reflections in Mathesius (1911) are particularly revealing in this respect. Mathesius observes that some linguistic phenomena, such as the length of English stressed vowels, are characterized by a range of variability defined by specific boundaries. This range of variability is associated

to the individual speaker: not only different speakers produce different values of the same phenomenon, but there exists a range of different realizations for one linguistic variable within the speech of the individual speaker. Therefore, in order to study the range of variation of a phenomenon, it is necessary to start from the inspection of the concrete realizations of a linguistic variable. According to this view, individualism is not equivalent to atomism or fragmentation, inasmuch as there are clear limitations to the range of variation.

The groundbreaking paper by Weinreich et al. (1968) represents a fracture with respect to the contemporary framework of linguistic studies; at the same time, it develops some central tenets of European linguistics at the beginning of the twentieth century. In fact, the authors devote much attention to the European traditions concerning the study of variation, and to Mathesius himself. More specifically, the programme of contemporary European linguists in the framework of individualism is revised according to typically American beliefs, by introducing an emphasis on the regularities inherent in variation and on the community pattern, rather than on the variability of the individual.

In European individualism, the relationship between variation and the limits of variation, as well as between heterogeneity and invariant structure, was discussed by making an appeal to the behavior of the individual speaker; the relationship between the individual and the community was problematic in itself. In particular, Mathesius was very clear about the dialectics between oscillation and limits of the oscillation. On the other hand, Weinreich, Labov and Herzog criticize Mathesius precisely for not having sought structured heterogeneity: the individual's behaviors were seen as referring to a class of speakers, also because of the influence of macro-sociological models. The specificity of the individual as historically determined source is reduced or nullified also because of the massive use of statistics, which brought to light regularities. In this respect, the dominant trend of Anglo-Saxon sociolinguistics (the so-called "first wave of social studies of variation", in Penelope Eckert's words; Eckert 2012) achieved interesting results on the analysis of diastatic variables, but these results did not appear to be revolutionary in the domain of stylistic variation.

Models of multidimensional variation such as those obtained, for example, for (r) in the North-American context have rarely been applied in the European context (apart from the English area): this result cannot be caused only by the different research areas of European scholars. The call for regularity and structured heterogeneity is consistent with the correlational research method, but above all it is related to the particular socio-cultural conditions of North-America. Therefore, a philological, historical study of spoken texts (where by 'historical' we mean related to the conditions of the text and of the speaker producing it) is desirable,

especially in Europe. The individual speaker is considered as a source and, consequently, all his/her characteristics should be documented: in this sense, data collection should be anthropological and microlinguistic.²

As far as the principle of microscopy is concerned, sound archives enable us to observe the opposition between 'local' and 'global' at its best: microscopic study allows the best correlation between the three dimensions of variation (diaphasy, diatopy, diastaty). Sometimes, it is precisely microscopy (the study of the oscillations determined by a certain phenomenon in the texts of a group of speakers of a community) which allows the understanding of linguistic change, while macroscopy helps us understand, on a large scale, the spreading factors external to the phenomenon. The difficulty lies, once again, in relating microscopic and macroscopic dimensions.

The principle of oscillation and that of linguistic context put two closely connected characteristics together under the same label. The oscillations in the realization of a certain phenomenon (in relation with Mathesius' concept of 'potentiality') directly recall the concept of 'variable' and its range of variants. As is well-known, linguistic context represents a fundamental source of variation: context-induced allophonic dispersion, for example, characterizes synchronic variability and variation and can prelude diachronic changes.

The principle of having recourse to spontaneous speech assumes that precisely at the spontaneous speech level the speaker selects variants which are in his/her competence, but do not come out in formal style or in response to direct questions. Direct elicitation methods, in fact, do not ensure the attainment of those levels of 'automatic' spontaneity which appear to be crucial when studying stylistic variation. Even within identical portions of spontaneous speech, differences in the level of self-control have been detected in which the speaker provides oscillations between Italianized and dialectal forms.

Several authors belonging to the European tradition have concentrated on the importance of psychological and linguistic differences between speakers. In Gilliéron, Dauzat, Gardette and Duraffour (e.g., Gillierón & Roques 1912; Gillieron 1918, 1921; Dauzat 1900, 1922; Duraffour 1932; Gardette 1983) it is possible to identify a French line of theoretical and methodological reflection which gives prominence to the psychological aspects of production and knowledge of patois (Sornicola 2002). This certainly constitutes a crucial problem, especially for sociophonetic

2. The archive thus has the individual speaker as its basic unit of analysis. The archiving procedure is carried out by the interviewer him/herself (in order to minimize the loss of information from the person who was present at the recorded event and the person who archives it). The interviewer also prepares a sheet for every speaker in which to note down both the objective features of traditional sociolinguistic research (sex, age, education, job) and the subjective features (personality, attitudes, motivations that emerged during the interview).

variation, and can be summarized in the following question: if factors like age, education, occupation, family culture, and group culture coincide, then why is it the case that speakers can, and do, exhibit a wide range of linguistic behaviors?

Different answers have been put forward – including the impact of such factors as social ambition or local culture loyalty. The social network model and the level of interaction between the speakers within the network have recently been brought in. Nevertheless, these answers cannot satisfy the wide range of variation inside a peer group. A sound archive centered on spontaneous speech and including adequately long texts allows the study of the levels of heterogeneity and oscillation in relation to the differences in the speaker's level of linguistic automatism and consciousness. Heterogeneity and textual oscillations are expressions of the high polymorphism characterizing many phenomena of the area, especially diphthongization, as will be shown in §3 and §4.

3. Phlegraean diphthongs

Sound materials are collected in the Phlegrean area, along the Neapolitan Coast (the area of Naples and the Gulf of Pozzuoli, together with the islands of Capri, Procida, Ischia): from a linguistic point of view, this area is intriguing for several reasons. First of all, it appears to be a single physical reality, characterized by territorial homogeneity. But the geographic criterion clearly does not suffice to determine the linguistic interest of a territory. Also history contributes to the uniqueness of this area, which is culturally less connected to Naples than to the areas of Vesuvius and Sorrento. Actually, the Phlegrean territories seem to have somewhat resisted the dynamics of 'neapolitization' affecting the whole Campania region and vast areas of Southern Italy. Speakers from the islands and Pozzuoli alternate complete adherence to 'Neapolitan' norms and features and absolute loyalty to local characteristics/traits. It should also be remembered that understanding the periphery helps understand the center; therefore, a better understanding of the Phlegrean area will help to understand the complex dialectal variability of Naples, a city extremely rich in linguistic variation, which has not been explored from the sociophonetic point of view.

The dialect of the Gulf of Pozzuoli, and those of the islands of Procida and Ischia, are characterized by a context-dependent alternation between monophthongs and diphthongs. This phenomenon is very pervasive, because it involves four vocalic variables – (i), (e), (o), (u) – and is not limited by syllable structure.³

3. The properties of the syllabic environment of the diphthongs of Pozzuoli have been recently studied by Abete (2011).

The discussion will be limited here to the behavior of mid-high vowels /e/ and /o/, both in open and in closed syllables, which diphthongize in /ai/ (/ei/) and /au/ (/ou/), respectively (cf. Freund 1933:9 and 12; Rohlfs 1966–1969, §62 and §80).

In Rohlfs' account of Upper Southern Italy diphthongization of mid-high vowels a one-to-one mapping of variant and place is assumed, as it is shown in Figures 2 and 3. This representation does not take into account the high polymorphism affecting each community.

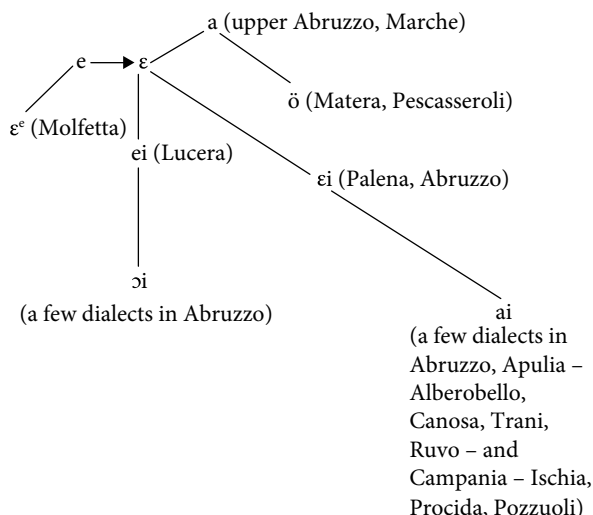


Figure 2. Geographical distribution of the variants of the mid-high front vowel (source: Rohlfs 1966–1969: 84–85).

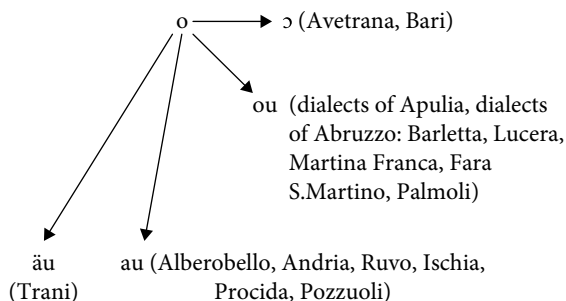


Figure 3. Geographical distribution of the variants of the mid-high back vowel (source: Rohlfs 1966–1969: 99).

Although a certain number of geographical variants of both vowels are reported for other places of upper southern Italy, Rohlf's transcribes as <ai> the main development from [e], and as <au> the main development from [o] in Pozzuoli, Ischia and Procida (see Table 1).

Table 1. Diphthongized variants from [e] and [o].

	open syllable	closed syllable
e > [ai]	<i>naivə</i> 'snow'; <i>vainə</i> 'vein'	<i>jənaistrə</i> 'broom'; <i>sajaittə</i> 'small boat'; <i>saikka</i> 'slim'; <i>aissa</i> 'she'
o > [au]	<i>vauçə</i> 'voice'; <i>nauçə</i> 'nut'; <i>çaurə</i> 'flower'; <i>saulə</i> 'alone'; <i>nəpautə</i> 'nephew/niece'	<i>saurdə</i> 'deaf (fem.)'; <i>raussə</i> 'red (fem.)'

Rohlf's recognizes that diphthongization influences vowels unaffected by the characteristic Neapolitan metaphony. For this reason the diphthongs we are dealing with have traditionally been called 'spontaneous' in the Italian dialectology literature. In some places, notably Pozzuoli and Forio d'Ischia, the 'spontaneous' diphthongal processes affect all vowels, except [a].⁴

The Phlegraean diphthongs pose both structural and dialectological questions. Structural aspects can be summarized as follows:

- a. high polymorphism (several vocalic variables are involved);
- b. syllable-independence;
- c. correlation to prosodic factors (a word such as *pisci* 'fishes' can be rendered as [pʏiʃi] before phrase boundaries, while it will be rendered as [piʃʃ] in internal position).

Dialectological questions can be described as follows: several scholars observed an apparent connection between the diphthongs of the Tyrrhenian coast located in the North of Naples (in particular, the islands of Ischia and Procida), and those occurring in the Adriatic region (Abruzzo, Molise, Apulia), which show a similar range of variants. Diphthongization of stressed mid vowels has been described for various places of Upper Southern Italy in the traditional dialectological literature. It was first pointed out by Salvioni (1911) that these processes have a peculiar geographical distribution on the Tyrrhenian and Adriatic sides of Italy (the Phlegraean area and Abruzzi and Apulia respectively), which led the Swiss scholar to identify a 'Tyrrhenian-Adriatic corridor' (for a discussion see Sornicola 2006a).

4. For Pozzuoli, a more accurate phonetic representation of stressed vowels, which also takes into account their developments from Latin and the systemic relationships between metaphonic and 'spontaneous' diphthongs, has recently been put forward by Abete (2011) and Abete & Simpson (2010).

Although the two areas show common features (e.g. a strong tendency to diphthongize stressed vowels, an evident instability of each diphthong), apparent differences also exist:

- a. Adriatic diphthongization is limited to open syllables, whereas Phlegraean diphthongization occurs regardless of syllable structures;
- b. Adriatic diphthongization is stress-sensitive (it is blocked when the word is a proparoxytone, and sometimes when it is an oxytone; Rohlf 1966–1969: 53–54), whereas Phlegraean diphthongization occurs regardless of stress pattern;
- c. Adriatic diphthongization involves all the vowels which are not affected by metaphony, whereas Phlegraean diphthongization is limited to the mid (and sometimes high) vowels.

To summarize, in the Adriatic area the process appears to be more regular and to have a more pervasive effect on the inventory, involving a high number of vowels (including the central-low vowel) and few contexts (open syllables, paroxytones). On the contrary, it involves a lower number of vowels and a wider range of contexts in most of the places of the Phlegraean area.

It is still unclear whether the Phlegraean diphthongs originated as spontaneous developments within the different speech communities or caused by population mixing. Undoubtedly, there is proof of various waves of migrations from the Adriatic coast and inner areas of Campania to the islands of Ischia and Procida. Ancient and modern historical sources attest that there was contact between the Phlegraean area and Apulia from ancient times up to the midpoint of the 18th century, because of traditional fishing activities in the area, such as the cultivation of oysters and mussels. Other sources attest demographic movements of non-exactly specified dimensions towards the Phlegraean islands from the inland regions of Campania and the coast of Abruzzo and Marche and even Romagna especially from the 16th and 17th centuries onwards. A study of the documents in the *Libri delle nascite, matrimoni e morti* in the Abbey of San Michele in Procida shows that in the period 1750–1799 34 individuals overall came to the island, 20 of whom were from the Adriatic coast; in the period 1800–1859 a total of 65 individuals came to the island, 29 of whom were from the Adriatic coast and the rest was from the inland regions of Campania. Significant Apulian immigration is registered in the *Libro dei Matrimoni* for the period 1873–1908: out of approximately 490 marriages, we find 85 in which at least one of the spouses (and in some cases both) is of Apulian origin (compared with 50 involving people from Gaeta and 27 from Sicilia). Trani, Alberobello, Andria, Monopoli and especially Molfetta are the cities of origin most frequently recorded (Sornicola 2006b). Nevertheless, this does not seem to provide sufficient evidence that in the Phlegraean dialects the diphthongs are not an indigenous development.

4. The diphthongal variability: Data from the archive

According to a well received view maintained by Rohlfs and other scholars, the Phlegraean diphthongs are rather recent, which could explain their inner instability and external variability. This theory seems far from convincing: diphthongal trajectories are inherently unstable and highly sensitive to context, therefore the possibility for diphthongs to keep polymorphism for centuries should not be surprising. The analysis of many stretches of spontaneous spoken discourse collected in ADICA and produced by different individuals from the Phlegraean area has shown that the diphthongs investigated have a conspicuous number of variants. Structural variation and intra-speaker variability phenomena are better treated according to a variationist methodology than to a historical-grammar approach.⁵

The shape of the diphthongs shows high variability across speakers. Let us concentrate on the variants of the front mid-high vowel. The overall ranges of variants of (e) in the Phlegraean dialects can be represented as follows (see Sornicola 2003 for further details): (e) = {e, ei, ε, εi, ε^l, εΛ, ε^Λ, æ, əi, ʌi, ʌⁱ, ʌ^ε, ʌ^ə, ʌ^v, ʌ^l}.

Strong individual variation has been detected in speakers of the island of Ischia. In a study conducted on three speakers from the village of Panza, close to the town of Forio (Sornicola 2001), significantly different ranges of variants and grammatical / lexical distributions emerged. Speaker I had a smaller and smoother range of variants which can be represented as follows: (e) = {e, ei, ε, æ, ʌ, ʌ^ε}

The presence of these variants in grammatical and lexical contexts is rather uniform, as shown in Table 2.

Table 2. Grammatical and lexical contexts.

variants	imperfect indicative forms	nominal forms
[e]	[tə'nevə] 'I had'	[rum'menəkə] 'Sunday'
[ε]	[e ssən'tevə] 'I herd them'	–
[æ]	[nu bbu'lævənə] 'they did not want'	–
[ʌ]	[pu'tʌvə] 'I could'; [e ssən'tʌvə] 'he herd them'; [kanuʃ'ʃʌvə] 'I knew'	
[ʌ ^ε]	[arra'pʌ ^ε və] 'I opened'; [nu bbu'lʌ ^ε vənə] 'they did not want'; [ʃkri'vʌ ^ε vənə] 'they wrote' (2 tokens); [və'nʌ ^ε və] 'I came'	[lum'mʌ ^ε nəkə] 'Sunday'
[e ⁱ]	–	[mul'le'irəmə] 'my wife'

5. Rohlfs, however, was well aware of the multiple outcomes of vowel diphthongization, but tended to attribute them to different diachronic phases and different geographic points.

Speaker II has a polarized range of variants with uneven grammatical and lexical distribution: (e) = {e, ε, ʌ, ʌ^e}.

While there is no diphthongization in nominal contexts, verbal forms and pronominal contexts show the phenomenon, as can be seen in Table 3.

Table 3. Grammatical contexts.

variants	imperfect indicative forms	pronominal context
[e]	[tə'nevə] 'I had'	
[ε]	[rum'menəkə] 'Sunday'	–
[ʌ]	[tə'nʌvənə] 'they held'; [fa'ʌvə] 'I did'; [vu'lʌvə] 's/he wanted'	[ʼmmʌ] 'me'
[ʌ ^e]	[nnə kɕju'lʌ ^e vənə] 'they closed us'; [te'nʌ ^e vənə] 'they held'; [pu'tʌ ^e vənə] 'they could'; [və'nʌ ^e və] 's/he came'	[ʼkʌ ^e ʝə] 'that (fem.)'

Speaker III too has a polarized range of variants, but unlike speaker II, this is rich in lowered and back forms: (e) = {e, eⁱ, ʌ, ʌ^e, ʌ^ə, ʌ^ʊ}. Examples are given in Table 4.

Table 4. Grammatical and lexical contexts.

variants	imperfect indicative forms	nominal forms	pronouns
[e]	[tə'nevə] 'I had'	[rum'menəkə] 'Sunday'	[ʼkeʝə] 'that (fem.)'
[e ⁱ]	[appartə'ne ⁱ vənə] 'they belonged'	–	–
[ʌ]	[ssə nə 'ʌvə] 's/he went away'; [vu'lʌvə] 's/he wanted'; [e rreflət'tʌvə] 'they reflected them'; [a'vʌvə fə] 's/he had to do'; [kanuʃ'ʌvə] 's/he knew'; [tra'sʌvə] 's/he went in'	[kur'tʌʝə] 'court', [ʼlʌttərə] 'letter'	[a m'mʌ] 'to me', [ʼʌssə] 's/he'
[ʌ ^e]	[tə'nʌ ^e və] 'I had, s/he had'; [tə'nʌ ^e vənə] 'they held'; [skri'vʌ ^e və] 'I wrote', [fa'ʌ ^e vənə] 'they did'	[pja'tʃʌ ^e rə] 'pleasure', [lum'mʌ ^e nəkə] 'Sunday'	[a m'mʌ ^e] 'to me'
[ʌ ^ə]	[tə'nʌ ^ə vənə] 'they held'	–	–
[ʌ ^ʊ]	[e ssəʃ'ʌ ^ʊ və] 's/he brought up'	–	–

Interestingly, despite the above mentioned differences, all three speakers always have the variants [ʌ], [ʌ^e], [ʌ^ʊ] in apocopated infinitive. This seems to be crucial evidence of the importance of the prosodic components of stress as factors influencing the diphthongal processes investigated. To the same conclusion points the fact that – whatever the individual and areal variability – in Pozzuoli and Forio, triphthongs are possible when the vowel has unusually high pitch/loudness/duration (i.e. in conditions of heavy stress): [di'çəʌⁱvə] 's/he said', [ʼsəʌʊrdə] 'deaf (fem.)'.

The influence of syntactic and prosodic factors on vowel movements has been repeatedly observed for numerous languages. As for the first domain, a correlation between diphthongization and [+Focus] syntactic positions has been reported in the literature from various languages. As for the second, pre-pausal lengthening appears to correlate with diphthongal movements: the final position in a phrase seems to be the most favorable context for diphthongized variants. In the case of Phlegraean [e], there seems to be a correlation between duration, pitch, loudness and diphthongization, while there does not seem to be a strong correlation between diphthongization and syntactic position associated to focus.

The data collected in Ischia show an extremely composite picture (Sornicola 2001, 2002). There is a non-negligible number of contexts in which – according to the expectations – the vowels in non-pre-pausal words (typically finite verbal forms that precede an object noun phrase) are not diphthongized; furthermore, a certain number of contexts show the diphthongized vowel [e] in pre-pausal words that have [+Focus] feature, as shown by the variants of the verbal form *tenere* ‘to have’ in Table 5.

Table 5. Alternation between monothongs and diphthongs in Ischia.

subject	no diphthongization	diphthongization
<i>speaker II</i>	[nui kə ttə'nevan-a 'terrə // tə'nevanə a 'rrɔbba / ma ki n-tə'nevə 'njendə] ‘those of us who had lands had possessions, but those who did not have anything...’	[al'lorə 'datəsə kə nnujə a 'rrɔbba a tə'nɫʷənə // nunn-əmpur'tavə] ‘then as we had possessions we were not interested in that’
<i>speaker III</i>	['kistə 'litfə kə 'ttenə kan'dine 'lwogi e 'ffwokə / te'nəvə a kan'dinə / te'nəvənə pprop'rjetə] ‘this (man) says that he has cellars, and other riches [literally ‘places and fires’], he had a cellar, they had properties’	[u 'ssalə ŋ'gapa ənd-a kuku'ttɛλλa o ttə'nɫʷə] ‘he was a sensible person’ (literally ‘the salt in his head, he got it’)

However, there are also contexts at odds with expectations. In (1) the utterance produced by the speaker is a sequence of three intonational phrases (the first two are noun phrases, the third is a temporal clause) with the semantic function of enumerating the times when the speaker drank wine. The vowel [ɛ] in the first phrase is diphthongized, while the two [e] in the second and the third are not, although no significant difference in stress among these vowels can be detected and the three syntactic and intonational units are separated from one other by a long pause:

- (1) Interviewer: [kwann-'erə na 'fɛstə?] ‘(you used to drink wine moderately) when there was a celebration?’
- Speaker I: [na 'fɛɫstə // na rum'mənəkə // 'kwannə və'nevən-a'miʃə] ‘(at) a celebration, on Sunday, when friends used to come’

In (2) both occurrences of the word *lummeneka* ‘Sunday’ have a diphthongized [e] ([lum'mΛ^enəkə]), although only the second token can be analyzed as a pre-pausal and [+Focus] noun phrase, while the first is a pre-pausal, circumstantial unit which does not have the [+Focus] feature, for it is clearly an afterthought uttered with the intonation contour of backgrounded information.

- (2) Speaker I: [kjam'ma o tə'lefənə iə / o 'mjeləkə // a lum'mΛ^enəkə] ‘it was me who phoned the doctor, on Sunday’
 Interviewer: [pek'ke / era e rum'menəkə?] ‘why, was it on Sunday?’
 Speaker I: [e: // er-e lum'mΛ^enəkə] ‘yes, it was on Sunday’

Finally, in (3) both intonational groups (two short clauses, the second being the repetition of the first) have strong prosodic Focus on the pre-pausal verbal form and are followed by a long pause. Yet only the second group has a diphthongized vowel in the pre-pausal verbal form.

- (3) Speaker III: [pa'pa vu'leva // pa'pa vu'lΛva] ‘daddy wanted (it), daddy wanted (it)’

In Pozzuoli too the situation is rather composite. On the one hand, the variant [Λⁱ] often appears in isolated and emphatic words (i.e. at the end of a tonal group), while the variant [Λ] often appears in pre-focal position in verbal forms (see (4) and (5)).

- (4) [kə ffa'fΛⁱvə?] ‘what did he do?’
 (5) [kə ppa'pa di'fΛⁱva / iə vu'lessə kam'pa pə vvə'rΛ a 'finə r a 'gwerrə] ‘that daddy was saying: I would like to live to see the war end’

On the other hand, in the same speakers [Λⁱ] does not appear only in focal contexts, and the variant [εⁱ], which otherwise appears frequently in non-focal contexts, can occur at the end of a tonal or focal group (see (6) and (7)).

- (6) ['kesta 'si a fa'fεⁱvənə] ‘this, yes, they did it’
 (7) [si nun tə'nΛⁱv a pa'ʃjentsja] ‘if he wasn't patient’

Other speakers from the same community show an even less advanced stage of diphthongization, in which the variants [εΛ] and [ε^Λ] appear in focal contexts (see (8)).

- (8) [nu bbu'le^Λvənə / nu bbu'le^Λvənə / nu bbu'le^Λvənə e nə'ssuna ma'njera] ‘they didn't want, they didn't want, they didn't want at all’.

5. Building sound archives to study linguistic variation

The high variability of the diphthongization of Plegrean /e/ can be compared with analogous phenomena occurring in different Romance dialects. As stated in §2, the Phlegrean area is located on the border of an extended area whose most relevant feature is the liveliness of spontaneous diphthongization, with particular regard to mid vowels. In the Romance-speaking domain, this area includes the Adriatic coast from Romagna to Abruzzo and Apulia, but also Franco-provençal dialects, the Retho-romance domain, together with the Dalmatian dialect (Sornicola 2003). Spontaneous diphthongization appears also in Tuscany (in Leghorn, but also in Arezzo; Calamai 2004, 2012). Experimental studies are still in their early stages and they do not usually consider uncontrolled speech styles. Yet the focus on controlled speech styles runs the risk of missing highly variable phenomena such as those treated in the present study.

A dialect sound archive focused on spontaneous speech, in addition to showing its potentials for a diatopic and therefore strictly dialectological analysis, undoubtedly has a sociolinguistic, and especially diaphasic, value. It allows the study of certain variables which, at first, show a high degree of instability. Phenomena like spontaneous diphthongization usually occur below the speaker's level of consciousness, are the product of spontaneous speech, do not show any correlation with classical sociolinguistic parameters and clearly represent a problem for the traditional description of the grammar of a dialect. The social distribution of this phenomenon lacks a clear pattern of variation in terms of age, sex, and social class. Strong individual variability has been detected in various places of the area: as demonstrated in §3, diphthongs irregularly occur in the spontaneous speech of a large part of the population. The sociolinguistic role of professional groups is unclear, although fishermen stand out as the social group that most regularly produces them. Precisely for these reasons, Phlegrean diphthongs call for a stylistic analysis. The study of the differences in the levels of production and consciousness and of the sometimes apparently unjustified changes in style, which can be offered by a spontaneous speech archive, is essential not only for understanding stylistic variation, but also for variation in a more general sense. The differences in the levels of production and consciousness are precious instruments for a hermeneutics of texts and speakers. This is a crucial point for sociolinguistics moving back towards dialectology: how can we, from a hermeneutic point of view, use the differences between the speakers, and their microhistories, to investigate intra-speaker variation? The hesitations in the choice or production of certain variants, and the changes in motivation or attitudes causing intratextual variation, have always been at the core of Romance dialectological tradition. Suffice it to think of the work notes of the *Atlante Italo-Svizzero* (Jaberg & Jud 1928–1949),

which were extremely illuminating in this respect: despite being associated with the traditional questionnaire technique, they tried to register all the complexity of the speaker's behavior through multiple answers. Spontaneous speech is actually the place of multiple answers, very fertile ground for the analysis of phonetic variation patterns: usage data allow the inspection of the variants, followed by attempts of generalization which take into account such factors as prosodic organization of the utterance, organization of turns, sociolinguistic functions. A similar model, *mutatis mutandis*, is employed by Temple in this volume: phonetic detail, however minute it may be, can convey multiple meanings on different levels.

In addition to representing a precious resource for stylistic analysis, sound archives also allow us to introduce the analysis of linguistic phenomena to a truly historical perspective. The growing attention to the Intangible Cultural Heritage which is spreading in different parts of the world may offer the scientific community a huge quantity of audio recordings by anthropologists, dialectologists, ethnographers from the entire twentieth century (Ginouves 2011). It thus becomes possible to envisage a historical experimental sociophonetics, and therefore to "repeat the past" by means of the analysis of old recordings, which represent "invaluable data for the study of change in the community, and for the studies of change or the absence of change in individual systems" (Labov 1994: 77). The limits of these real-time comparisons are evident, but certainly counterbalanced by the information on sound change which can be obtained through an intelligent exploitation of these sound resources, which have only rarely been used so far.

6. Conclusion

By disposing of the large quantity of spontaneous speech gathered in sound archives, we may rely on a solid empirical basis for the analysis and description of phenomena that do not appear to be related to classical sociolinguistic variables. This possibility is particularly important for those phenomena that pertain to the dialectal substratum / dialectal competence of the speakers, that is, to varieties that are hardly evocable by many techniques of speech elicitation, including the most sophisticated ones (such as map-tasks), and certainly impossible to be captured through traditional questionnaires. The so-called spontaneous diphthongization attested in the Phlegrean varieties as well as in several Romance areas is one of these problematic phenomena, being diphthongization captured only with difficulty by traditional dialectological and sociolinguistic analyses. The many sound archives collected throughout the Peninsula turn out to be a goldmine for socio-phonetic variation, and for the most part they are still unexplored.

References

- Abete, Giovanni. 2011. *I processi di dittongazione nei dialetti dell'Italia meridionale. Un approccio sperimentale*. Roma: Aracne.
- Abete, Giovanni & Adrian Simpson. 2010. "Confini prosodici e variazione segmentale. Analisi acustica dell'alternanza monottongo/dittongo in alcuni dialetti dell'Italia meridionale". *La dimensione temporale del parlato – Atti del V Convegno Nazionale dell'Associazione Italiana di Scienze della Voce (AISV)* ed. by Stephan Schmid, Michael Schwarzenbach & Dieter Studer, 297–323. Torriana: EDK.
- Calamai, Silvia. 2004. *Il vocalismo tonico pisano e livornese. Aspetti storici, percettivi, acustici*. Alessandria: Edizione dell'Orso.
- Calamai, Silvia. 2011. "Ordinare archivi sonori: il progetto Gra.fo". *Rivista Italiana di Dialettologia* 35. 135–164.
- Calamai, Silvia. 2012. "Il parlato aretino, tra dialettologia e analisi fonetica sperimentale". *Laboratorio critico* 2/2. 1–20. <http://ojs.uniroma1.it/index.php/laboratoriocritico>
- Calamai, Silvia & Pier Marco Bertinetto. 2012. "Per il recupero della Carta dei Dialetti Italiani". *Coesistenza linguistiche nell'Italia pre- e postunitaria* ed. by Tullio Telmon, Gianmario Raimondi & Luisa Revelli, 335–356. Roma, Bulzoni.
- Dauzat, Albert. 1900. *Études linguistiques sur la Basse-Auvergne: Morphologie du patois de Vinzelles*. Paris: Bouillon.
- Dauzat, Albert. 1922. *La géographie linguistique, avec 7 figures dans le texte*. Paris: Flammarion (2nd edition: 1948).
- Di Salvo, Margherita. 2006. "Digitalizzazione e catalogazione di tre corpora di parlato". *Bollettino Linguistico Campano* 9–10. 1–282.
- Duraffour, Antonin. 1932. *Phénomènes généraux d'évolution phonétique dans les dialectes franco-provençaux, d'après le parler de Vaux-en-Bugey (Ain)*. Grenoble: Institut phonétique de Grenoble.
- Eckert, Penelope. 2012. "Three waves of variation study: The emergence of meaning in the study of variation". *Annual Review of Anthropology* 41. 87–100.
DOI: 10.1146/annurev-anthro-092611-145828
- Freund, Ilse. 1933. *Beiträge zur Mundart von Ischia*. PhD Thesis, Tübingen, Eberhard-Karls-Universität.
- Gardette, Pierre. 1983. *Études de géographie linguistique*, publiées par les soins de Brigitte Horiot, Marie-Rose Simoni, Georges Straka. Strasbourg: Société de Linguistique Romane.
- Gillieron, Jules. 1918. *Généalogie des mots qui désignent l'abeille, d'après l'Atlas linguistique de la France*. Paris: Champion.
- Gillieron, Jules. 1921. *Pathologie et thérapeutique verbales*. Paris: Champion.
- Gillieron, Jules & Mario Roques. 1912. *Études de géographie linguistique d'après l'Atlas linguistique de la France*. Paris: Champion.
- Ginouès, Véronique. 2011. "Quand le renard raconte ses histoires au monde. La naissance du portail du patrimoine oral, catalogue collectif d'archives sonores et audiovisuelles". *Internationale de l'Imaginaire (Le patrimoine culturel immatériel: premières expériences en France)* 25. 107–128. <http://halshs.archives-ouvertes.fr/halshs-00588487/fr/>
- Jaberg, Karl & Jakob Jud. 1928–1940. *Sprach- und Sachatlas Italiens und der Südschweiz, Vol. I–VIII*. Zofingen: Ringier.
- Labov, William. 1994. *Principles of Linguistic Change, Vol. I: Linguistic Factors*. Oxford: Blackwell.

- Mathesius, Vilém. 1911. "On the potentiality of phenomena of language." *A Prague School Reader in Linguistics* ed. by Josef Vachek, 1–32. Bloomington & London: Indiana University Press.
- Rohlf, Gerhard. 1966–1969. *Grammatica storica della lingua italiana e dei suoi dialetti*, Vol. I *Fonetica*, Einaudi, Turin.
- Salvioni, Carlo. 1911. "Zur Lautgeschichte. Appunti per la storia del vocalismo tonico italiano." *Zeitschrift für Romanische Philologie* 35. 486–488.
- Sornicola, Rosanna. 2001. "Alcune recenti ricerche sul parlato: le dinamiche vocaliche di (e) nell'area flegrea e le loro implicazioni per una teoria della variazione". *Scritto e Parlo. Metodo, testi e contesti. Atti del Colloquio Internazionale di Studi* ed. by Maurizio Dardano, Adriana Pelo & Antonella Stefinlongo, 239–264. Roma: Aracne.
- Sornicola, Rosanna. 2002. "La variazione dialettale nell'area costiera napoletana. Il progetto di un archivio di testi dialettali parlati". *Bollettino Linguistico Campano* 1. 131–155.
- Sornicola, Rosanna. 2003. "Polimorfismo e instabilità strutturale: un esame della dittongazione spontanea dell'area flegrea in una prospettiva romanza". *Actas del XXIII Congreso Internacional de Lingüística y Filología Románica* ed. by Fernando Sánchez-Miret, 301–313. Tübingen: Max Niemeyer.
- Sornicola, Rosanna. 2006a. "Dialectology and history. The problem of the Adriatic-Tyrrhenian dialect corridor". *Rethinking Languages in Contact. The Case of Italian* ed. by Anna Laura Lepschy & Arturo Tosi, 127–145. Oxford: Legenda.
- Sornicola, Rosanna. 2006b. "Dialecto e processi di italianizzazione in un habitat del Sud d'Italia". *Lingua e dialetto nell'Italia del duemila* ed. by Alberto A. Sobrero & Annarita Miglietta, 195–242. Galatina: Congedo.
- Weinreich, Uriel, William Labov & Marvin I. Herzog. 1968. "Empirical foundations for a theory of language change". *Directions for historical linguistics* ed. by Winfred P. Lehmann and Yakov Malkiel, 95–188. Austin: University of Texas Press.

PART III

What is (and what is not) a sociophonetic change

Ejectives in English and German

Linguistic, sociophonetic, interactional,
epiphenomenal?

Adrian Simpson

University of Jena

This paper describes the phonetic form, the distribution and the possible functions of ejectives in English and German, proposing that ejectives are on the increase in different varieties in English. The problems of teasing apart the different contributions of allophonic regularity, interactional function, sociophonetic variability and epiphenomenal inevitability in accounting for ejectives in English are discussed. Possible production mechanisms behind ejectives in both languages are explored and doubt is cast on previous epiphenomenal accounts which have ignored the importance of a pulmonic component in creating the necessary intra-oral pressure increase. This, in turn, raises questions about possible production mechanisms behind ejectives in languages in which they play a regular part in the phonological inventory.

1. Introduction

Ejectives are a well-documented and much analysed sound class in the languages of the world. Conventionally, ejective consonants are described as being the product of an airstream mechanism that comprises a closed glottis and the vertical, upward movement of the larynx (e.g. Catford 1977) producing either an increase in intraoral pressure for a plosive burst or sufficient airflow for a fricative. Indeed, behind the term ‘ejective’ lie a number of different, if related, production mechanisms in different languages (Kingston 1985, 2005; Ladefoged & Maddieson 1996; Lindau 1982, 1984; Warner 1996; Wright et al. 2002).

Until only recently ejectives in European languages, such as English and German, have generally only found sporadic and brief mention in the research literature with little systematic description of their form, function or distribution (Gordeeva & Scobbie 2006). This is despite the fact that, in different varieties of

British English at least, ejectives would seem to be on the increase. In this paper, I will draw together the observations in the literature on ejectives in English and in German, two languages in which ejectives are considered to play at most only a very marginal role. I will examine the problems we confront when trying to describe the functions they are carrying out and how this could be inextricably linked to the ways in which they might be being produced. In particular, the form and elusive distribution of ejectives in English raise a number of interesting challenges for sociophonetic analysis. For this reason, rather than being quantitative, this paper is essentially qualitative and, in places, speculative as to questions of distribution and, indeed as to how ejectives might actually be being produced. However, central to this paper, as it is to others in this volume (Stuart et al., Temple), is the time that I will spend looking at and at times casting serious doubt on what it is we are actually describing.

2. Ejectives in English

Although this needs to be empirically verified, ejectives would appear to be becoming more frequent in many varieties of English. It is possible that this has to do with changing expectations or changes in analytical and observational techniques, but it seems unlikely that the widespread occurrence of ejectives in a range of English varieties today would have escaped the attention of such an acute observer as Catford, who mentions the occurrence of ejectives in English only in passing: “in English they occasionally occur as the realization of final [p, t, k] in pathological speech, and in some northern English dialects” (Catford 1977:70).

The apparent marginality is supported by the absence of any mention of ejectives in earlier detailed studies of (pre-)glottalisation in English (Higginbottom 1964; Roach 1973, 1979). Although later, Roach (2002: 24) states that “[i]n English we find ejectives allophones of /p, t, k/ in some accents of the Midlands and North of England”. Wells (1982), however, does attribute the use of ejectives to “both northerners and southerners” (1982:261).

It is also possible that an increase in the prevalence of ejectives in English is merely a further development in the increase in the prevalence of pre-glottalised plosives and glottal replacement which Roach (1973, 1979) was primarily concerned with, although Collins & Mees (1996) show us that we need to be cautious when jumping to conclusions about the timeline of apparent changes.

2.1 Structural distribution

All studies are in agreement that ejectives in English occur finally, although further details are less clear. So, while Catford (1977:70) merely states that ejectives can occur as realisations of final /p, t, k/, Roach (2002:24) restricts the context to “the end of a word preceding a pause”. The most detailed attempt to describe the structural contexts in which ejectives can occur is undertaken in Ogden (2009), based primarily on data from naturally occurring Scottish English:

Ejectives occur:

- word finally (and not e.g. before vowels)
- in stressed syllables
- after vowels, nasals and laterals (which are all voiced), but not after voiceless sounds such as [s]
- within utterances (before pauses), as well as at the end of utterances.

(Ogden 2009: 163)

2.2 Sources and functions

It is one thing to recognize that ejectives occur in a language, another to describe what functions they are fulfilling in the language in question. Ejectives could arise from at least four different sources. In English it is possible that they are simply allophonic, i.e. a contextual or conditioned variant of plosives. This would imply that in at least certain varieties of English, just as it is possible to predict a pulmonically fuelled aspirated plosive in the onset of a stressed syllable, so we should be able to predict the occurrence of an ejective in a particular final context. This does not seem to be the case. So, while Ogden (2009) provides a number of contexts in which ejectives can occur, the presence of the context itself does not guarantee the occurrence of an ejective. However, epiphenomenal ejectives in German and possibly some of those in English might be candidates for this type of prediction (see below).

More plausibly, and as a possible prior stage to a more systematic phonological status, ejectives in many varieties of English are part of the sociophonetic variation employed by individual speakers in different communicative contexts. To date little sociophonetic analysis of ejectives has been undertaken, but one study supports this analysis for Scottish English. Gordeeva & Scobbie (2006) analysed the speech of seven pre-school children and showed that approximately 10% of word-final stops were realized as ejectives, although no more direct prediction of the occurrence than this is provided.

Ogden (2009) describes and illustrates tokens of ejectives using examples from naturally occurring talk from a Scottish English speaker. Ogden's examples illustrate well the different interactional functions ejectives might be fulfilling. In common with other descriptions, ejectives occur finally in his data, but the structural intricacy of conversational data allows for a more detailed analysis. So, for instance, ejectives seem to be one of the correlates of floor-holding pauses, first described in Local & Kelly (1986) in a data set initially described by Jefferson (1983). Such a floor-holding pause was found to begin with glottal closure and end with its release. The ejective shown in the example in (1), taken from Ogden (2009) would seem to be part of such a pausal complex.

- (1) *at- [ʔ] on the week]end of wee[k' (0.3 s) ʔ] three* (Ogden 2009: 165)

In (1) the closed glottis extends from the end of the word *week* to the beginning of the word *three*. It would be possible to interpret the ejective release of the dorsal plosive as an explicit expression of the glottal closure.

These two analyses of ejectives from Gordeeva & Scobbie (2006) and Ogden (2009) highlight a possible conflict between two theoretically different interpretations of the same data set. In Ogden's approach, a range of different phonetic patterns are accounted for in terms of the work they are doing in structuring conversation, and variation per se plays a subordinate role. By contrast, while Gordeeva & Scobbie's analysis does take account of structure (e.g. finality), a more detailed categorization in terms of interactional structure is not present – the concentration is on ejectives as part of the possible set of patterns of variation. However, analysing sociophonetic variation using only restricted contextual information, such as finality, is dangerous. Assigning different phonetic shapes to the same sociophonetic variable assumes that information about different and identical structural context is known. But as Example (1) illustrates, within normal interaction, there are different categories of word-finality or pre-pausality, different contexts which may or may not be accompanied by different bundles of phonetic events. Another well-documented example of this is the high prevalence of word-final plosive aspiration found in Tyneside English speakers producing word-list material (Docherty et al. 1997; Local 2003). A possible sociophonetic interpretation of this, but one which only considers word- or utterance-finality as a structural categorization, is that the higher frequency of occurrence of aspirated release is an approximation to standard forms. However, in an earlier study on the phonetic shape of turn-taking in Tyneside English, Local et al. (1986) found that the same phonetic pattern, i.e. the aspirated release of voiceless plosives, is one of the phonetic correlates of turn-finality. It is hardly surprising, therefore, that speakers of this variety of English should produce aspirated final plosives in word-lists, producing the phonetics of turn-finality after each word (Local 2003).

Indeed, the situation is complicated further by our failure to know exactly which of her/his rich set of phonetic resources a speaker is bringing to bear on a situation when speech is being produced outside the context of naturally occurring talk. The phonetic patterns that speakers have at their command are undoubtedly put to their most systematic and most diversified use in the course of normal conversational interaction. By contrast, other activities, such as reading aloud word lists or texts put artificial demands on speakers, leading us to expect phonetic patterns at apparently the same structural place being variable because any mapping from interactional structures to those of reading aloud texts or word lists must be ambiguous and lead to varying degrees of unsystematic and arbitrary transfer into such an artificial situation.

The variable presence of ejectives in English both at an inter- as well as at an intraindividual level presents us with a similar set of problems. The suspected increase in the prevalence of ejectives in different varieties of English over the last few decades must correlate with patterns of sociophonetic variability involving ejectives that were not previously present. At the same time, it is still unclear what the structural contexts are from a linguistic or interactional point of view. Data drawn from two non-spontaneous sources make this clear and at the same time emphasise the ambiguity described above. The first data set is the first series of the television comedy *The Office*. From approximately three hours of material a total of eight ejectives were identified. Spectrograms and oscillograms of three examples shown in (2) from three of the main characters are shown in Figure 1. The braces in (2) indicate the extent of the excerpts shown in the Figure 1.

- (2) a. *it's not often you get something that's {both <outbreath> romanti[k'] and thrifty}*
 b. *he's perf[k'] (1.0)*
 c. *and whether (0.1 s) they can (0.35 s) {pay (0.45 s) for i[t']} (1.75 s)*

In line with our description so far, all of these examples are word-final, but (2a) is by no means pre-pausal as the plosive releases directly into the vowel of *and*. And in the remaining two pre-pausal examples there is no sign of these pauses being turn-holding. Indeed, we can hypothesise that turn-holding is a feature absent in read speech, or at best will arbitrarily surface in theatrical dialogue. It has also been suggested that due to the burst intensity (see Figure 1a) ejectives correlate with emphasis (Wells 1982: 261) or enhance the consonantal place of articulation (Ogden 2009: 164).

A further data set appears, on the surface at least, to present a clearer case of sociophonetic variability in an otherwise uniform structure. Simpson (1992) describes the “glottal piece” in the naturally occurring talk of one speaker of Suffolk English. Put simply, the glottal piece represents a cooccurrence restriction

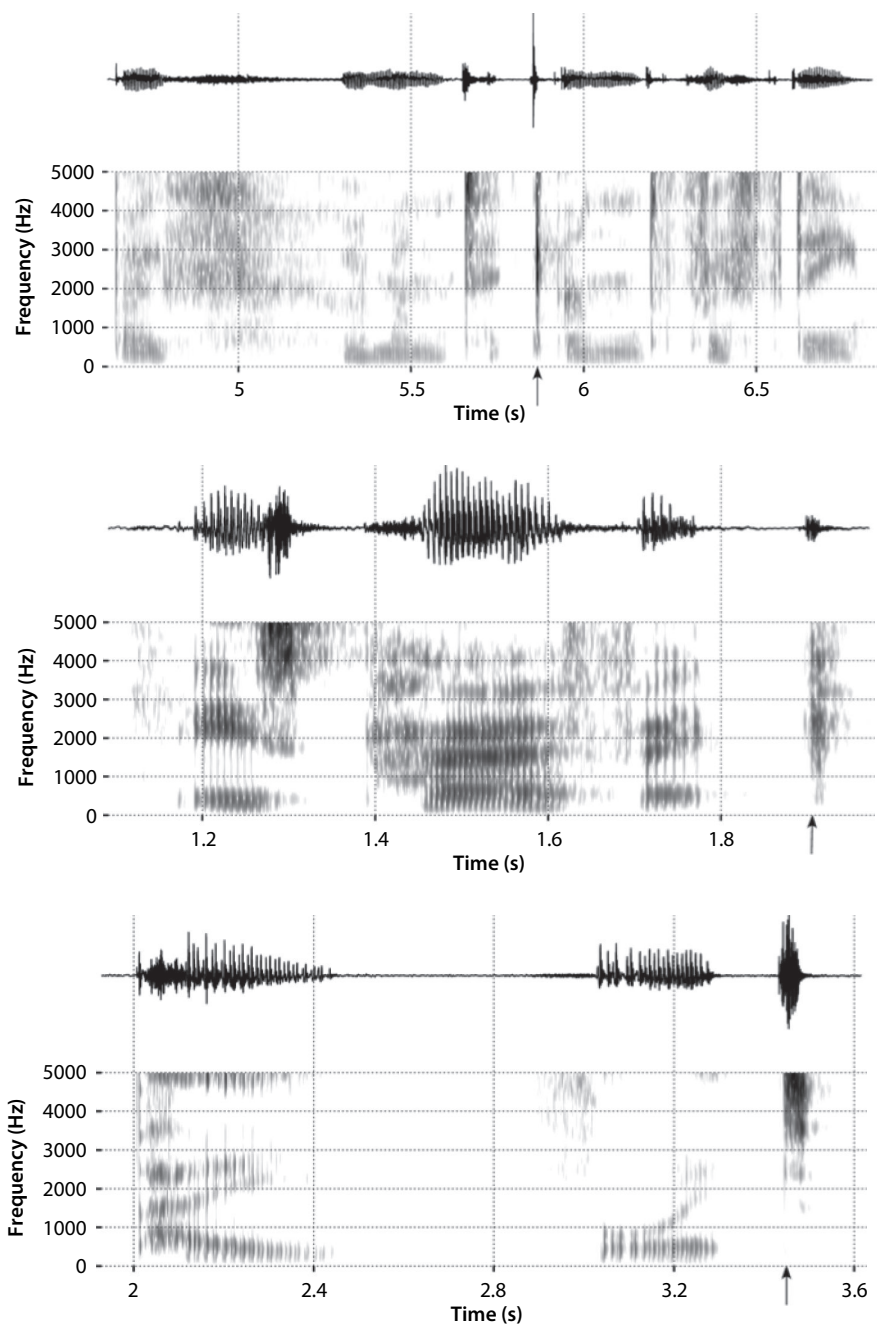


Figure 1. Oscillograms and spectrograms of ejective tokens from the TV comedy series *The Office*. Vertical arrows indicate the location of the ejective in each example: (a) *romantic*, (b) *perfick* and (c) *it*.

on the presence of two glottal stops either side of [ə], e.g. [pʊʔə] *put a* but [pʊɾəʔ] *put it* (see also Trudgill 1974; Lodge 1984). In an attempt to investigate this phenomenon instrumentally using electroglottography (EGG), the same speaker was recorded producing a list of short sentences. Although not of direct relevance at the time of the experiment, the words *back* or *work* had been placed at the end of several of the sentences. Despite the lexical and syntactic simplicity of the sentences (e.g. *She'll look at it at work*), the speaker exhibited a number of dysfluencies (pauses, false starts, truncations) when producing many of the sentences. This was most likely due to a considerable discrepancy between her naturally occurring patterns and the standard-like pronunciation the speaker considered appropriate to the formality of the recording situation. However, of direct interest to the discussion here was the realisation of the final plosives in 'back' and 'work'. The plosives were always released, but in approximately one third of cases the plosive was realised as an ejective. Figure 2a shows an example of an ejective release, Figure 2b a pulmonically fuelled release. The oscillogram shows the EGG trace, which I will return to below. Approaching these patterns from a sociophonetic point of view, it is possible to treat the ejectives as being one part of the standard-like patterns being produced by this speaker in the context of a formal recording situation, perhaps again being a correlate of articulatory place enhancement, as Ogden (2009) has suggested. But even here, it is possible to propose that, although the sentences may represent structural identity from the analyst's point of view, there is no guarantee that the speaker is treating them this way. So, even if our speaker is producing a sequence of sentences in a studio setting, it is not legitimate to state that observable differences in the realisation of final plosives in a series of sentences

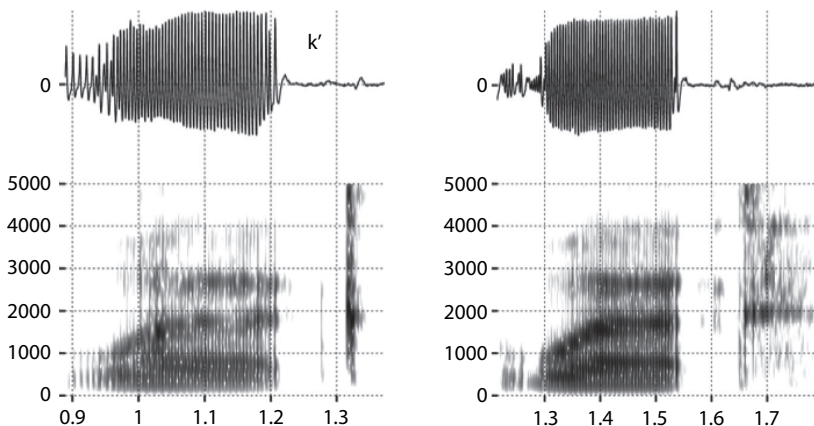


Figure 2. EGG and spectrograms of (a) ejective vs. (b) pulmonically released plosives from sentence-final tokens of the word *back* from a Suffolk English speaker.

with final ‘work’ are necessarily structurally or contextually the same when she is producing them. In other words, she may be making recourse to different interactional strategies when producing what superficially look like two reproductions of the same expression, producing the phonetic details of floor-holding, say, in one case, the phonetic details appropriate to yielding the floor in another.

3. Epiphenomenal ejectives and production mechanisms

The conventional way of describing the production mechanism behind ejectives assumes a closed glottis, a raised velum, a supraglottal stricture of complete closure (plosives) or close approximation (fricatives) and an upward movement of the larynx causing the supraglottal air pressure to rise. Catford’s (1977: 79) schematic representation of this mechanism is shown in Figure 3.

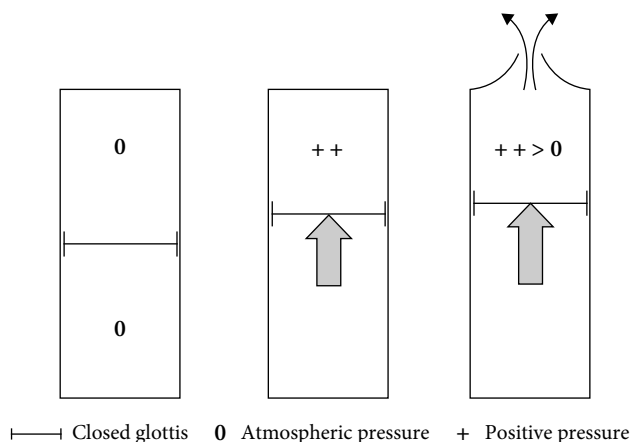


Figure 3. Schematic representation of the production mechanism behind an ejective (“glottalic pressure unphonated”) from Catford (1977: 79).

What might the production mechanism(s) be that are fuelling ejectives we have been describing in English? The articulatory and phonatory suggestions made in the works cited generally fail to provide details about exactly how ejectives are being produced. Ogden (2009: 164) suggests that they may “involve a rearrangement in time of the constrictions needed to produce glottally reinforced voiceless plosives.” However, Wells (1982: 261) states that “[a]n emphatic articulation of the glottal component will readily convert this into an ejective” without providing further details. It is likely that both Ogden and Wells are assuming the accepted method of ejective production outlined above. But doubt has been cast on this

account. In particular, Kingston (1985) shows using a digital implementation of Rothenberg's aerodynamic model (Rothenberg 1968; Müller & Brown 1980) that larynx raising alone is not able to produce the necessary increase in supraglottal pressure to produce the types of ejective bursts he observes in languages such as Tigrinya. Instead, he suggests that additional articulators may be employed to reduce supraglottal cavity size, such as tongue root backing.

However, in an attempt to find an adequate account of the production mechanism for epiphenomenal ejectives in German and as a consequence also ejectives in English, we must even cast doubt on the initiatory contribution of decreasing the size of the supraglottal cavity to increase intraoral pressure and fuel an ejective plosive release.

In German, epiphenomenal ejectives are the product of a temporal overlap of a final plosive and junctural glottalisation, i.e. glottal stop or creak at vowel onset. Figure 4 compares (a) an epiphenomenal ejective release from *weht ein* 'blows a' with (b) the pulmonically fuelled release of a final pre-pausal plosive in *mit* 'with' from Simpson (2007). This paper looks at different ways in which both nasal and oral stops in German can be produced with releases fuelled by non-pulmonic airstream mechanisms. So, for instance, in nasal-plosive sequences such as that in *in Kiel*, we routinely find a weak, yet consistently present, click at the point of the apical release of the nasal. This click is an epiphenomenon produced following a double, apical-velar closure followed by the release of the frontmost (apical) closure. A sufficient change in air pressure needed to give rise to the click is produced by a small change in the size of the intraoral cavity prior to apical release (Ohala 1995). Likewise, in line with Ohala (1997), Simpson (2007) suggests that in the ejective release of plosives we find in an example such as that shown in Figure 4a, a change in supraglottal air pressure is brought about by vowel-to-vowel movement taking place during the double glottal and oral closure. However, there are two things that might lead us to question this interpretation. First, pressure change due to vowel-to-vowel movement predicts that we should find plosive releases fuelled by both glottalic egressive (ejectives) as well as ingressive (voiceless implosives). In the first case, intraoral pressure increases due to a vowel-to-vowel movement bringing about a reduction in the size of the supraglottal cavity, e.g. [a-i]. In the second case, intraoral pressure decreases due to an enlarging vowel-vowel movement, such as [i-a]. However, no voiceless implosive releases were found. Secondly, impressionistically, many such plosive releases seem to be intenser than might be expected from a lot of vowel-to-vowel movements, e.g. [e] to [a] in *weht ein*. Although these doubts are based primarily on auditory impression and visual interpretation of the acoustic record, it does seem worth speculating about other production mechanisms behind such ejectives.

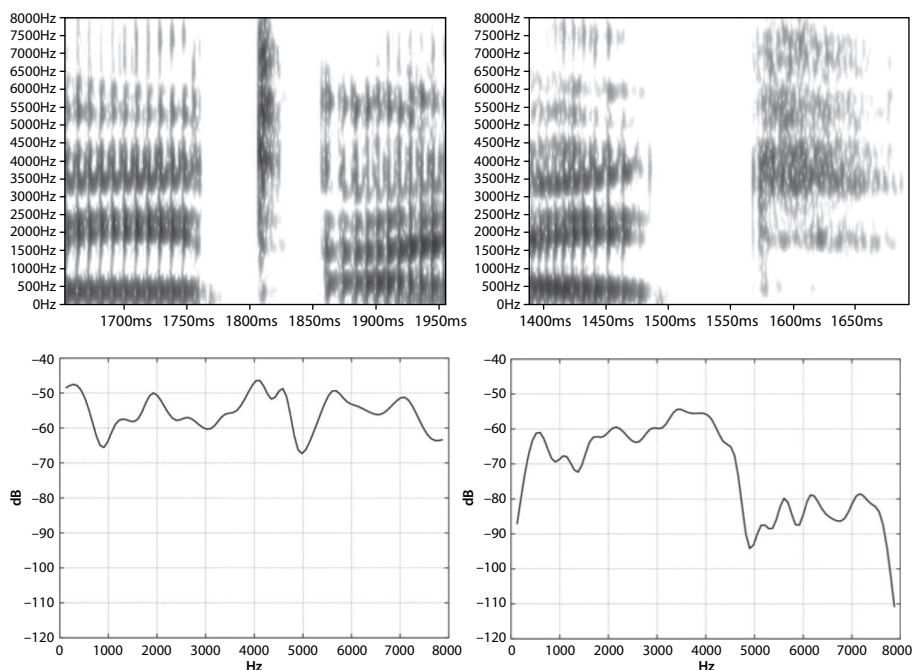


Figure 4. Spectrograms and DFT sections of (a) ejective release of word-final plosive in *weht* (*ein*) 'blows' and (b) pulmonically released word-final and pre-pausal plosive in *mit* 'with'. Vertical arrows indicate location of plosive release (from Simpson 2007).

The first account is that there is active movement of the larynx involved. However, this account is problematic as it implies that such ejectives are not epiphenomenal, but actively produced, and the question of motivation arises, i.e. why would a speaker choose to actively raise the larynx in this context? An enhancement of the place of the articulation of the plosive in a burst which might otherwise be masked by glottal closure would be a possible motivation. Nevertheless, a second account seems more plausible since, on the one hand it explains the relative intensity of some release bursts, and at the same time the explanation remains epiphenomenal. This account assumes that air is still flowing through the glottis after the supraglottal closure has been made, but prior to the point of oral stop release the glottis is closed or configured for creak. The suggested mechanism is shown in Figure 5 using Catford's (1977) method of schematisation. Following supraglottal stop closure, intraoral air pressure increases due to pulmonic airflow through the glottis. At some point prior to the release of supraglottal closure, the glottis closes or is configured for creak. When the oral closure is finally released, the burst is auditorily and acoustically an ejective, although the pressure build-up

is pulmonically fuelled and no active movement of the larynx or any other articulator has taken place. This account is still only informed speculation and awaits confirmation from the results of intraoral pressure measurements in combination with transillumination of the glottis which are planned beyond this study.

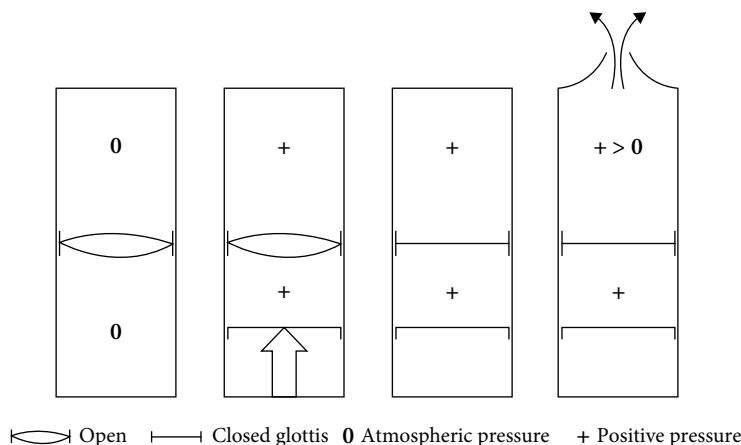


Figure 5. Representation of production mechanism driving an epiphenomenal ejective release using Catford's (1977) schematisation method.

4. English ejectives revisited

Finding an epiphenomenal account of ejectives in German that does not involve active movement of the larynx throws a different light on how ejectives in English and other languages might be being produced. It might also lead us to expect that ejectives fulfilling different functions in the same language may be the product of different mechanisms.

Ogden's (2009) suggestion that the temporal realignment of the articulatory and phonatory components of glottally reinforced plosives fits well with the account of epiphenomenal ejectives in German since it would imply that glottal closure is synchronised with plosive release. However, what is not made explicit in Ogden's account is exactly how intraoral pressure build-up would occur. This is however, accounted for if we assume that during the first part of stop closure, the glottis would still be open allowing an initial pulmonic airstream to flow into the supraglottal cavity.

Two further observations of patterns in English and German support this account. The first observation is negative: data from both languages have yet to

provide any evidence of ejective fricatives. The epiphenomenal account predicts this. It implies a pulmonically driven intraoral pressure build-up prior to glottal closure and finally oral release of stop closure. A similar sequence of glottal events synchronised with the stricture of close approximation of a fricative will produce pulmonically fuelled friction followed by its cessation once glottal closure is made. Interestingly, final ejective fricatives would also appear to be absent in English, as well (Ogden 2009).

A second piece of evidence suggesting a pulmonically fuelled pressure build-up is the data from Suffolk English presented above. In Figure 2 the spectrogram is aligned with the EGG signal. What one might have expected from the EGG trace is evidence of differences between the two plosive types *prior* to stop closure. However, this was not the case. Glottal activity observed from the EGG signal showed no obvious differences for the ejective or the aspirated stop releases. Again, if we assume that there is a pulmonically driven build-up of intraoral pressure during the initial phase of the plosive, the only difference between the two plosives would reside in whether the glottis was closed or open on release.

Although it is outside the remit of this paper, it is worth considering to what extent languages with ejectives as part of their regular phonological systems may also employ a similar mechanism. Indeed, in Rothenberg's aerodynamic model employed by Kingston (1985) to model various aspects of ejective production that was mentioned earlier, one set of model configurations employs a constant glottal flow to raise sufficient pressure for the ejective burst, i.e. it does not rely on the reduction in the size of the supraglottal cavity alone to produce the required change in air pressure. Furthermore, the presence of perseverative voicing well into the stop closure phase of all plosive types including ejectives in Georgian (Wysocki 2004; Grawunder et al. 2010) implies that at least part of the pressure build-up in some ejectives is pulmonically fuelled.

5. Discussion

Unlike epiphenomenal ejectives in German, ejectives are still an elusive feature in many varieties of English and for many speakers, as is evidenced by their relative paucity in the data set from the first series of *The Office*. Despite being able to coarsely describe some of the linguistic and interactional contexts in which they occur, predicting their occurrence for a particular speaker in a particular context is far from straightforward, and it remains a challenge to see whether it will be possible to provide a chronologically ordered sociophonetic analysis of the regional spread of ejectives through English. At present this is a daunting task. Despite there being several acoustic databases of English covering different

speakers of both sexes from different regions and social backgrounds spanning a number of decades, even annotated databases do not contain an adequate level of transcriptional detail to reliably analyse such features as the details of stop release. Annotated databases on the whole use a limited set of phonetic-phonological labels which allow the analyst to temporally locate tokens of particular linguistic and phonological categories, but provide unsystematic phonetic information, extremely coarse with respect to some details, more differentiated with regard to others. The annotations of the *Kiel Corpora of Read and Spontaneous Speech* (Simpson et al. 1997) provide good examples of this. Due to theoretical interests at the time of creating the labeled database, aspects of junctural glottalisation and glottalised reflexes of fortis and lenis plosives were recorded in some detail. However, other systematic details, such as stop releases fuelled by velaric or glottalic airstream mechanisms are merely subsumed under a general annotation of release. Similarly, annotated databases of English, such as the phonetically labeled sections of the British National Corpus (Coleman et al. 2011), IViE (Grabe et al. 1998) or DiVyS (Nolan et al. 2009), do not consistently contain direct information about ejectives, although given the spectral characteristics of ejectives (see above), a semi-automatic identification of annotated plosive releases should be possible.

Nevertheless, watching the development of ejectives in English is of particular sociophonetic interest. Sociophonetic variation often involves gradual changes that can be plotted along a single dimension, e.g. changes in vowel quality (e.g. Labov et al. 2006) or consonantal stricture (e.g. Cravens & Giannelli 1995). Alternatively, variation may arise from the import of a sound from another variety (e.g. Milroy et al. 1994). On the surface, ejectives in English do not fit neatly into either of these categories. Although now apparently present in many different varieties of British English at least, it is not clear whether any of these varieties can be seen as a source. It seems more likely that an internal source is responsible, one which perhaps involves one of the possible outcomes of the different temporal alignment of the glottal and articulatory components of (pre-)glottalised plosives. It has therefore been an important part of this paper to discuss possible production mechanisms behind ejectives in English, working backwards from considerations about the highly predictable epiphenomenal ejectives in German. Presumably, we must predict that speakers will produce ejectives in English in at least two different ways, one in which pulmonic airflow leads to a build-up of intraoral pressure, the other in which true glottalic initiation is used, using larynx-raising to compress the air trapped in the supraglottal cavity. From the point of view of analysing how sound patterns are perceived, reinterpreted phonatorily and articulatorily, and propagated by speakers throughout a community, ejectives represent an intriguing case. Instead of proposing a misinterpretation of the acoustic patterns, as Ohala has proposed as a possible source of certain sound changes (e.g. Ohala 1974, 1979;

Ohala & Busà 1995), the glottal and articulatory components which can produce potentially identical acoustic patterns and auditory impressions of an 'ejective' can be genuinely ambiguous. And indeed, there is no reason why there should not be intra- as well as interspeaker variability in the production of ejectives performing different functions.

Much of what has been said about the possible production mechanisms behind ejectives has been speculation informed by the acoustic record and impressionistic observation and remains to be substantiated by instrumental investigation, in particular using a combination of transillumination and air pressure measurement to examine and compare ejective production in a range of languages and different linguistic and interactional contexts.

Finally, I have kept away from any discussion of using whether 'ejective' is an appropriate term, whether it would be appropriate to use modifiers such as 'stiff' or 'slack', or whether some of the examples we have looked at should be treated as complex segment types. Instead, I have tried to concentrate on examining possible production mechanisms and functions of elements which are unified by giving rise to similar auditory impressions and acoustic records.

References

- Catford, John C. 1977. *Fundamental Problems in Phonetics*. Edinburgh: Edinburgh University Press.
- Coleman, John, Mark Liberman, Greg Kochanski, Lou Burnard & Jiahong Yuan. 2011. "Mining a Year of Speech". *VLPS 2011: New Tools and Methods for Very-Large-Scale Phonetics Research*, Philadelphia, University of Pennsylvania, 29th–31st January 2011.
- Collins, Beverley & Inger Mees. 1996. "Spreading everywhere? How recent a phenomenon is glottalisation in Received Pronunciation?". *English World-Wide* 17/2. 175–187.
DOI: 10.1075/eww.17.2.03col
- Cravens, Thomas D. & Luciano Giannelli. 1995. "Relative salience of gender and class in a situation of multiple competing norms". *Language Variation and Change* 7. 261–285.
DOI: 10.1017/S0954394500001010
- Docherty, Gerard J., James Milroy, Lesley Milroy & David Walshaw. 1997. "Descriptive adequacy in phonology: A variationist perspective". *Journal of Linguistics* 33. 275–310.
DOI: 10.1017/S002222679700649X
- Foulkes, Paul, James M. Scobbie & Dominic Watt. 2010. "Sociophonetics". *Handbook of Phonetic Sciences* ed. by William Hardcastle, John Laver & Fiona Gibbon (2nd edition), 703–754. Oxford: Blackwell. DOI: 10.1002/9781444317251.ch19
- Grawunder, Sven, Adrian P. Simpson & Madzhid Khalilov. 2010. "Phonetic characteristics of ejectives – samples from Caucasian languages". *Turbulent sounds. An interdisciplinary guide* ed. by Susanne Fuchs, Martine Toda & Marzena Żygis, 209–244. Berlin: De Gruyter Mouton. DOI: 10.1515/9783110226584.209

- Higginbottom, Eleanor. 1964. "Glottal reinforcement in English". *Transactions of the Philological Society* 63. 129–142. DOI: 10.1111/j.1467-968X.1964.tb01010.x
- Jefferson, Gail. 1983. "On a failed hypothesis: 'Conjunctionals' as overlap- vulnerable". *Tilburg Papers in Language and Literature* 28. 29–33.
- Kingston, John C. 1985. *The phonetics and phonology of the timing of oral and glottal events*. PhD Thesis, Berkeley, University of California.
- Kingston, John C. 2005. "The phonetics of Athabaskan tonogenesis". Athabaskan prosody ed. by Sharon Hargus & Keren Rice, 137–184. Amsterdam & Philadelphia: John Benjamins.
- Labov, William, Sharon Ash & Charles Boberg. 2006. *The Atlas of North American English*. Berlin & New York: Mouton De Gruyter.
- Lindau, Mona. 1982. "Phonetic differences in glottalic consonants". *Journal of the Acoustical Society of America* 71/S1. 22. DOI: 10.1121/1.2019283
- Lindau, Mona. 1984. "Phonetic differences in glottalic consonants". *Journal of Phonetics* 12. 147–155.
- Local, John K. 2003. Variable domains and variable relevance: interpreting phonetic exponents. *Journal of Phonetics* 31. 321–339. DOI: 10.1016/S0095-4470(03)00045-7
- Local, John K. & John Kelly. 1986. "Projection and 'silences': notes on phonetic detail and conversational structure". *Human Studies* 9. 185–204. DOI: 10.1007/BF00148126
- Local, John K., John Kelly & William H. G. Wells. 1986. "Some phonetic aspects of turn-delimitation in the speech of Urban Tynesiders". *Journal of Linguistics* 22. 411–437. DOI: 10.1017/S0022226700010859
- Lodge, Ken R. 1984. *Studies in the Phonology of Colloquial English*. London: Croom Helm.
- Milroy, James, Lesley Milroy, Sue Hartley & David Walshaw. 1994. "Glottal stops and Tyneside glottalization: Competing patterns of variation and change in British English". *Language Variation and Change* 6. 327–357. DOI: 10.1017/S095439450000171X
- Müller, Eric & William S. Brown. 1980. "Variations in the supraglottal air pressure waveform and their articulatory interpretation". *Speech and language: Advances in basic research and practice* 4. 317–389.
- Grabe, Esther, Francis Nolan & Kimberley J. Farrar. 1998. "IViE – A comparative transcription system for Intonational Variation in English". *Proceedings of the 5th International Conference on Spoken Language Processing* ed. by Robert H. Mannell & Jordi REobert-Ribes, 1259–1262. Australian Speech Science and Technology Association, Incorporated (ASSTA).
- Nolan, Francis, Kirsty McDougall, Gea de Jong & Toby Hudson. 2009. "The DyViS database: style-controlled recordings of 100 homogeneous speakers for forensic phonetic research". *International Journal of Speech Language and the Law* 16. 31–57.
- Ogden, Richard. 2009. *An introduction to English phonetics*. Edinburgh, Edinburgh University Press.
- Ohala, John J. 1974. "Experimental historical phonology". *Historical linguistics II. Theory and description in phonology. Proceedings of the 1st International Conference on Historical Linguistics (Edinburgh, September 2–7, 1973)* ed. by John M. Anderson & Charles Jones, 353–389. Amsterdam: North Holland.
- Ohala, John J. 1979. "The contribution of acoustic phonetics to phonology". *Frontiers of speech communication research* ed. by Björn Lindblom & Sven Öhman, 355–363. London: Academic Press.
- Ohala, John J. 1995. "A probable case of clicks influencing the sound pattern of some European languages". *Phonetica* 52. 160–170. DOI: 10.1159/000262167

- Ohala, John J. 1997. "Emergent stops". *Proceedings of the 4th Seoul International Conference on Linguistics (SICOL 1997)*, Seoul, 11th–15th August 1997. 45–50.
- Ohala, John J. & Maria Grazia Busà. 1995. "Nasal loss before voiceless fricatives: a perceptually-based sound change". *Rivista di Linguistica* 7. 125–144.
- Roach, Paul. 1973. "Glottalization of English /p/, /t/, /k/ and /tʃ/ – a re-examination". *Journal of the International Phonetic Association* 3. 10–21. DOI: 10.1017/S0025100300000633
- Roach, Paul. 1979. "Laryngeal-oral coarticulation in glottalized English plosives". *Journal of the International Phonetic Association* 9. 2–6. DOI: 10.1017/S0025100300001857
- Roach, Paul. 2002. A little encyclopedia of phonetics. <http://www.freewebs.com/raulillox/Enciclopedia.pdf>
- Rothenberg, Martin. 1968. "The breath-stream dynamics of simple-released-plosive production". *Bibliotheca Phonetica* No 6. Basel & New York: Karger. http://www.rothenberg.org/Breath-Stream/BSD_contents.htm
- Simpson, Adrian P. 1992. "Casual speech rules and what the phonology of connected speech might really be like". *Linguistics* 30. 535–548. DOI: 10.1515/ling.1992.30.3.535
- Simpson, Adrian P. 2007. "Acoustic and auditory correlates of non-pulmonic sound production in German". *Journal of the International Phonetic Association* 37. 173–182. DOI: 10.1017/S0025100307002927
- Simpson, Adrian P., Klaus J. Kohler & Tobias Rettstadt, eds. 1997. *The Kiel Corpus of Read/Spontaneous Speech – Acoustic data base, processing tools and analysis results* (= AIPUK, 32). Kiel: University of Kiel.
- Trudgill, Paul. 1974. *The Social Differentiation of English in Norwich*. Cambridge: Cambridge University Press.
- Warner, Natasha. 1996. "Acoustic characteristics of ejectives in Ingush". *Proceedings of the International Conference on Spoken Language Processing (ICSLP 1996)*, Philadelphia, 3rd–6th October 1996. New York: Institute of Electrical and Electronics Engineer. 1525–1528. DOI: 10.1109/ICSLP.1996.607907
- Wells, John C. 1982. *Accents of English, Vol.1: An Introduction*. Cambridge, Cambridge University Press.
- Wright, Richard, Sharon Hargus & Katharine Davis. 2002. "On the categorization of ejectives: data from Witsuwit'en". *Journal of the International Phonetic Association* 32. 43–77. DOI: 10.1017/S0025100302000142
- Wysocki, Tamra M. 2004. *Acoustic analysis of Georgian stop consonants and stop clusters*. PhD Thesis, Chicago, University of Chicago.

Authors index

- A**
- Abete, Giovanni 175, 177
 Agostiniani, Luciano 143, 146, 148
 Aitken, Adam J. 60
 Albert, Abraham A. 145
 Aristotle 90
 Ashby, William 41
 Ashton, Lindsay 85
- B**
- Baayen, R. Harald 5
 Barlow, Michael 50
 Baroni, Maria Rosa 159
 Barry, Martin C. 119
 Bermúdez-Otero, Ricardo 97–99, 119, 128–130, 132
 Bernhard, Gerlad 159
 Bertinetto, Pier Marco 12, 142, 170
 Bickerton, Derek 21
 Biliotti, Francesca 159
 Bladon, R. Anthony 126
 Blondeau, Hélène 24–25
 Boberg, Charles 28, 98
 Boersma, Paul 32
 Booij, Geert 41
 Borrelli, Doris 148
 Boyd, Sally 98
 Brillouin, Léon 38, 52
 Broadbent, Judith M. 118
 Browman, Catherine 121, 125–126, 129, 132
 Brown, W.S. 197
 Bucholtz, Mary 69
 Buizza, Emanuela 106
 Busà, Maria Grazia 202
 Bybee, Joan L. 1, 4, 32, 35, 39–41, 50–51, 129
- C**
- Calamai, Silvia 6, 8, 10–11, 98, 156, 159, 169–170, 183
 Calderone, Basilio 4
 Carey, Emma 73, 75
 Carlson, Rolf 139, 162
 Catford, John C. 189–191, 196, 198–199
 Celata, Chiara 4, 6, 98, 149
 Chevrot, Jean-Pierre 35, 50, 52
 Chomsky, Noam 31
 Clauset, Aaron 38
 Clédât, Léon 34
 Coetzee, Andrees 98
 Coleman, John 97, 201
 Collins, Beverley 190
 Coquillon, Annelise 35
 Coulmas, Florian 140
 Coupland, Nikolas 69
 Cravens, Thomas 159, 201
 Crowdy, Steve 101
 Cruttenden, Alan 108–109, 111, 119, 121–122
- D**
- Dauzat, Albert 174
 De Jong, Daan 34, 41
 Delattre, Pierre C. 84
 Dell, François 34
 Denes, Peter B. 88
 Deshaies-Lafontaine, Denise 138
 Di Salvo, Margherita 171
 Dixon, James M. 61
 Docherty, Gerard J. 32, 71, 89, 98, 115–116, 126, 138, 162, 164, 192
 Dobson, Eric J. 83
 Douglas, Fiona M. 65
 Duraffour, Antonin 174
- Durand, Jacques 4–5, 33–35, 41, 60
 Durkheim, Emile 18
- E**
- Eckert, Penelope 69, 87, 173
 Ellis, Lucy 119, 128–130
 Encrevé, Pierre 34, 36, 53
 Evans, Betsy E. 6
- F**
- Fabricius, Anne 116
 Fanciullo, Franco 148
 Filipponio, Lorenzo 160
 Fougeron, Cécile 34
 Foulkes, Paul 2, 32–33, 51, 71, 89–90, 98, 115–116, 126, 138, 162
 Freeman, Donald C. 84
 Freund, Ilse 176
 Fujimura, Osamu 82
- G**
- Galli De' Paratesi, Nora 159
 Gardette, Pierre 174
 Giannelli, Luciano 141, 143–144, 146, 159, 201
 Gili Fivela, Barbara 150
 Gillieron, Jules 174
 Gimson, Alfred Charles 108–109, 111, 121–122
 Ginouvès, Véronique 10, 184
 Goldinger, Stephen D. 1, 4, 32, 89–90
 Goldstein, Louis M. 121, 125–126, 129, 132
 Gordeeva, Olga B. 189, 191–192
 Grabe, Esther 201
 Grammont, Maurice 50
 Grant, William 60–61

- Grassi, Corrado 151
 Grawunder, Sven 200
 Gunter, Barrie 68
 Guy, Gregory 98, 128, 130
- H**
 Hall, Stuart 68
 Hamann, Silke 32
 Hardcastle, William J. 119, 121, 128–130
 Harris, Roy 90
 Hawkins, Sarah 89, 139, 162
 Hay, Jennifer 1, 70, 138
 Hayes, Bruce 105, 115
 Hazen, Kirk 98
 Henderson, Jannette B. 102–103
 Herman, Gabriel 90
 Herzog, Marvin I. 3, 9, 173
 Heselwood, Barry 74, 79, 88
 Higginbottom, Eleanor 190
 Holmes, Janet 18
 Holtus, Gunter 149
 Hoole, Phil 111, 119, 129, 131
 Hornsby, David 35
 Horvath, Barbara M. 158
 Horvath, Ronald J. 158
- J**
 Jaberg, Karl 183
 Jaffe, Alexandra 69
 Jannedy, Stephanie 1, 138
 Jauriberry, Thomas 66, 73, 91
 Jefferson, Gail 192
 Johnson, Daniel Ezra 22–24
 Johnson, Keith 32, 69, 82, 87, 89–90, 162
 Johnston, Paul 60–63, 65, 71, 73–74, 87
 Jud, Jakob 183
- K**
 Kaufman, Terrence 140
 Kelly, John 192
 Kemmer, Suzanne 50
 Kerswill, Paul 7, 20–21, 88
 Kingston, John C. 189, 197, 200
 Kiparsky, Paul 99
 Köpke, Barbara 3
 Kühnert, Barbara 111, 119, 129, 131
 Kuhl, Patricia K. 69
- L**
 Laberge, Suzanne 21–22
 Labov, William 3–4, 7, 9–10, 25–26, 32, 34, 64, 69, 89–90, 137–140, 150, 156, 158, 163, 173, 184, 201
 Laks, Bernard 1, 4, 31, 34–35
 Langacker, Ronald W. 32, 40, 50
 Lausberg, Heinrich 8
 Laver, John 93
 Lawson, Eleanor 5, 61, 64–65, 71, 74–75, 79, 81–83, 85–87
 Lennon, Robert 79, 86
 Le Page, Robert B. 18
 Lepschy, Anna L. 142
 Lepschy, Giulio 142
 Lichtman, Karen 130
 Lindau, Mona 189
 Llamas, Carmen 66, 70, 73
 Local, John K. 139, 150, 192
 Lodge, Ken R. 131, 195
 Loporcaro, Michele 142–143, 148
 Lyche, Chantal 35
- M**
 Macafee, Caroline 62, 67
 MacFarlane, Andrew E. 60, 73
 Manuel, S. Y. 124
 Malécot, André 41
 Mallet, Géraldine-M. 35
 Marotta, Giovanna 8, 88, 143–144, 146–147, 149, 151–153, 155, 159–161
 Mathesius, Vilém 9, 172–174
 McClelland, James L. 40, 50
 Mees, Inger 190
 Milroy, James 158, 201
 Milroy, Lesley 158, 201
 Müller, E. 197
 Myers, James 99
- N**
 Nocchi, Nadia 144, 151, 160
 Nolan, Francis J. 7, 100–102, 105, 108, 111–112, 114, 116–117, 119, 122, 125, 128, 132, 201
- O**
 Ogden, Richard 71, 79, 88, 191–193, 195–196, 199–200
 Ohala, John J. 11, 88, 197, 201–202
- P**
 Pacini, Beatrice 146, 159
 Pagliano, Claudine 35
 Pater, Joe 98
 Payne, Arvilla 19–20
 Peirce, Charles S. 139, 140
 Pierrehumbert, Janet 1, 69, 89, 162
 Pinson, Elliot N. 88
 Plug, Leendert 71, 74, 79, 88
 Pukli, Monika 66, 73, 91
 Pulgram, Ernst 50
- R**
 Ranson, Diana 41
 Repp, Bruno H. 102–103
 Ricci, Irene 159
 Richmond, Suzanne 61
 Roach, Paul 190–191
 Roberts, Julie 98
 Robins, Robert H. 130
 Rohlfs, Gerhard 9, 152, 167, 176–179
 Romaine, Suzanne 61–63, 66, 71, 82
 Roques, Mario 174
 Rothenberg, Martin 197, 200
- S**
 Salvioni, Carlo 9, 177
 Sankoff, David 109
 Sankoff, Gillian 21–22, 24–25, 29, 97
 Santa Ana, Otto 98
 Savoia, Leonardo M. 143, 146, 152
 Scobbie, James M. 5, 62, 79, 98, 189, 191–192
 Schane, Sanford A. 32, 34
 Schuppler, Barbara 109
 Sebregts, Koen 79
 Silverstein, Michael 140
 Simpson, Adrian P. 10, 126–127, 131, 177, 193, 197–198, 201
 Smith, Jennifer 98
 Solé, Maria-Josep 126
 Sornicola, Rosanna 8–10, 33, 171, 174, 177–179, 181, 183
 Speitel, Hans-Henning 62, 71, 82
 Sproat, Richard 82
 Steriade, Donca 131–132

- Stevens, Kenneth 81
 Stuart-Smith, Jane 5–6, 60,
 62–63, 66–68, 71–74, 83,
 97–99, 116
- T**
 Tabouret-Keller, Andree
 18, 163
 Tagliamonte, Sali A. 7, 98–99,
 101, 104, 107, 109, 119, 126
 Telmon, Tullio 149
 Temple, Rosalind A.M. 6–7,
 97–99, 101, 104, 107, 109–110,
 116, 119, 126, 184, 190
 Timmins, Claire 63
 Thomason, Sarah G. 140
 Tollfree, Laura 158
- Tomasello, Michael 32
 Tranel, Bernard 34
 Trousdale, Gaeme 129
 Trudgill, Peter 66, 131, 140,
 156, 195
 Tucker, Benjamin V. 130–132
 Turchi, Laura 150
 Twist, Alina 87
- V**
 Volkart-Rey, Reinhardt 159
- W**
 Walker, Abby 132
 Warner, Natasha 130–131, 189
 Weinreich, Uriel 3, 9, 17–18,
 32, 173
- Wells, John C. 5, 60, 118, 158,
 190, 193, 196
 Wimmer, Gejza 38
 Williams, Anne 20–21, 61
 Williams, Irene F. 61
 Wolfram, Walt 98, 100, 102
 Wright, Richard 189
 Wright, Susan 7, 88
 Wysocki, Tamra M. 200
- Y**
 Yaeger-Dror, Malcah 71
- Z**
 Zhou, Xinhui 84, 87
 Zipf, George Kingsley 38

Subject index

- A**
- AAVE *see* African American Vernacular English
- accommodation 4
- adolescent 24, 62, 64–67, 69, 71
 middle-class adolescents 65
 working-class adolescents 64–66, 69
- affrication 8, 106, 143, 146–147, 149–150, 153–155, 157, 160–161
 s-affrication 8, 146, 149–150, 153–155, 157, 160–161
- African American Vernacular English (AAVE) 98, 102, 108, 129
- age 5, 9, 20–22, 24–27, 34–35, 41, 44–49, 51, 63, 71–72, 129, 151, 163, 174–175, 183
 age-grading 129
- airflow 11, 75, 189, 198, 201
- alveolar 41, 44, 73, 84, 102, 105–106, 108, 110–112, 115, 117, 120–121, 125–126, 128–130, 132, 143, 149
- ambition 160, 175
- apical 24–25, 60–61, 83, 85–86, 197
- apocope 143, 146–148, 153–155, 161, 167
- apparent time 9, 26
- approximant 60–63, 65, 71–76, 79, 83, 85, 87
 apical approximant 85
 bunched approximant 74
 postalveolar approximant 61–62, 73
 retroflex approximant 63, 83
 rhotic approximant 63
 uvular approximant 71–72, 75
- archive 9–10, 169–172, 174–175, 179, 183–184
 oral archive 10
 sound archive 8–10, 169–170, 172, 174–175, 183–184
 speech archive 183
- Articulatory Phonology 128–130
- assimilation 7, 99, 102, 105, 112, 118–122, 124–125, 128–129, 131, 143, 149
 anticipatory assimilation 125
 place assimilation 102, 105, 119
 progressive assimilation 124
 regressive assimilation 120
 voicing assimilation 102, 118, 125
- attitude 66–68, 155, 163, 174, 183
- attrition 3
- audio restoration 10
- auditory transcription *see* transcription
- awareness 9, 140–142, 149, 154, 157, 159, 163–164
- B**
- bai* 21–22
- bilingualism 3
- bunching 85–87
- burst 102–104, 189, 193, 197–198, 200
- C**
- caretakers 20–21, 24–25
- categoricity 119, 128–130, 132
- centralization 19
- change 2, 6, 8–11, 18, 22, 24–26, 59–60, 62, 64–71, 81, 86–88, 91–92, 114, 119, 129, 138, 141, 158–159, 161, 163, 174, 183–184, 190, 197, 200–201
 change from above 65
 change from below 62, 65
 contact-induced change 66, 69–70, 91
 vernacular change 62, 65–66
- child 2–3, 17, 19–26, 28, 35, 52, 61–62, 64, 73, 150, 191
- class 5, 9, 19, 25, 26–27, 44–45, 59–60, 62–66, 69, 71–74, 79, 83–88, 91, 151, 158–159, 172–173, 183
 middle-class 19, 25–26, 59–60, 62–63, 65, 71, 79, 83–84, 86–88
 lower working class 27, 62
 social class 5, 9, 26, 60, 65, 151, 159, 172, 183
 working class 25–27, 59–60, 62–63, 64–66, 69, 71–74, 79, 84–87, 91
- click 197
- closure 101–102, 109, 111–112, 115, 192, 120, 125–126, 132, 150, 159, 192, 197–200
 alveolar closure 125–126, 132
 apical-velar closure 197
 bilabial closure 111–112
 coronal closure 112
 glottal closure 115, 192, 198–200
 oral closure 197
 stop closure 102, 150, 159, 199, 200
 supraglottal closure 198

- cluster 97–105, 107–109, 112, 114–118, 121–122, 125–127, 130, 132
- word-final clusters 97–98, 100–101, 108, 122
- coalescence 102, 122–125
- coda 34, 36, 59–60, 62, 64, 66, 73–74, 82, 91, 99, 117, 130, 144, 149, 151–152
- cohesion 4, 35, 50–51
- internal cohesion 4, 35
- syntactic cohesion 50–51
- community 3–4, 6, 8, 17–19, 21–22, 24–25, 28, 32, 59–60, 68–70, 73–74, 86–89, 92, 138–141, 156–158, 161, 163–164, 173–174, 176, 178, 182, 184, 201
- linguistic community 4, 32, 139–140, 156
- social community 8, 140
- speech community 3, 17–18, 25, 28, 70, 86–87, 157, 161, 163–164
- compensation 87
- perceptual compensation 87
- competence 4, 32, 140, 153, 161, 174, 184
- compound 108, 115
- Connected Speech Processes (CSPs) 7, 97, 99–100, 125, 132, 135
- consonant 5, 7, 19, 32, 34–36, 41–45, 49–51, 60, 62, 66–69, 74–75, 81, 97–112, 114–122, 124–127, 129, 132, 142–144, 146–149, 151–153, 155, 159–161, 189
- consonant class 49, 155
- consonant
- strengthening 148
- consonant type 41–42
- Italian consonant system 142
- latent consonant 50
- linking consonant 35, 41
- conspiracy 160, 167
- contact 3, 4, 6, 10, 19, 28, 66–70, 86, 91, 167, 178
- dialect contact 67–69
- direct contact 68, 70
- indirect contact 68–70, 86
- continuum 6, 60, 62–63, 65, 70–74, 79, 84–85, 87, 89, 100–101, 105, 107, 112, 122, 129, 140
- rhotic-derhotic continuum 6, 63, 65, 70–71, 73–74, 79, 89
- social-gender continuum 62
- control 8, 24, 89, 127–128, 154–156, 159, 174
- conversation 10, 34, 36–37, 127, 152, 157, 192
- free/spontaneous conversation
- with peers 34
- guided one-to-one conversation 34
- see also* interaction
- cophonation 7, 125
- corpus 4–5, 7, 11, 31–33, 35–52, 59–60, 62–66, 71, 79, 81, 83, 85, 99, 101, 201
- corpus linguistics 12, 32–33
- creaky *see* voice quality
- creole 21
- CSPs *see* Connected Speech Processes
- culture 137, 159, 175
- family culture 175
- group culture 175
- cursus language 50
- D
- Dalmatian *see* dialect
- datum* 31–32
- de-affrication 146–147, 153–154
- deletion 6–7, 71, 97–99, 102, 107–112, 114, 116, 118, 121–122, 125–130, 143, 147–148, 152, 154, 159–160
- categorical deletion 116, 128
- coronal stops deletion 98
- derhoticisation 5–6, 59–71, 74–75, 78–79, 81–83, 85–86, 91
- detail 2, 10–11, 89, 97, 100, 102, 120, 126–128, 131, 139, 146, 150, 152, 159–170, 184, 196, 201
- phonetic detail *or*
- fine phonetic detail 2, 10, 98, 100, 102, 120, 126–128, 139, 146, 150, 152, 159, 170, 184, 196
- devoicing 99, 118, 122, 131
- dialect 8, 17, 19–20, 24–27, 32, 60, 62, 66–69, 73, 98, 108, 112, 116, 126, 138, 140, 142–144, 147–148, 151, 163, 167, 171, 175–176, 178–179, 183, 190
- Dalmatian dialect 183
- Florentine dialect 142
- Franco-provençal dialects 183
- Glaswegian dialect 62
- Italian Central dialects 148
- Italian Northern dialects 143
- Italian Southern dialects 143, 148
- New York City dialect 26
- Quebecois dialect 24
- southern US dialects 108
- Tuscan dialect 142
- York dialect 112
- see also* contact
- dialectology 2, 7, 138, 141, 160, 172, 177, 183
- historical dialectology 2
- traditional dialectology 7, 138, 160
- urban dialectology 172
- digital preservation 10
- diphthong 8–9, 61–62, 169–170, 175, 177–179, 181, 183
- diphthongal movements 181
- Phlegrean diphthongs 9, 169, 178, 183
- diphthongization 9, 178, 183–184
- Adriatic
- diphthongization 178
- spontaneous diphthongization 9, 183–184
- distribution 4–5, 20, 31–32, 35, 37–46, 48–49, 51–52, 63–64, 82, 97, 101, 103–104, 130, 143, 148, 169, 176–177, 179–180, 183, 189–191
- lexical distribution 143, 179–180
- Mandelbrot-Zipf distribution 38
- power-law distribution 4, 31, 35, 38–39
- duration 73–75, 109, 111, 180–181

- E**
 EAT *see* Edinburgh
 Articulation Test
 Edinburgh Articulation Test
 (EAT) 61, 92, 143, 153
 education 44, 47–49, 137,
 158–160, 163, 174–175
 education level 5, 35, 41, 44,
 46–49, 51, 160
 EGG *see* electroglottography
 egressive 197
 ejective 10, 11, 126, 189–202
 electroglottography (EGG)
 195, 200
 electromagnetic articulography
 (EMA) 111, 130
 EMA *see* electromagnetic
 articulography
 emphasis 4, 89, 173, 193
 emphatic word 182
 English 3, 5–7, 10–11, 25, 28,
 59–63, 65–71, 73, 75, 82–83,
 86–91, 97–99, 102, 108, 111,
 114, 116, 118–120, 124, 131–132,
 151, 158, 163, 172–173, 189–193,
 195–197, 199–201
 American English
 69, 87–88, 108, 131, 163
 British English 7, 10, 73, 97,
 99, 102, 111, 114, 116, 158,
 190, 201
 East Anglian 131
 English English 60, 66–67,
 69–70, 73, 83, 91
 Highland English 62
 London English 67–69,
 86, 168
 Scottish English 5, 59–63,
 67, 69–70, 82, 86–89,
 91–92, 191–192
 Standard Scottish English
 (SSE) 60, 62, 69
 Southern English 60, 66
 Suffolk English 131, 193,
 195, 200
 Tyneside English 192
 ethnicity 26, 28
 Jewish ethnicity 27
 Exemplar Theory 1, 4, 32, 89,
 97, 162
Exemplum 31
 explicitness 100–101, 106, 112,
 114, 122, 125–127, 131
- F**
 F1 *see* formant
 F2 *see* formant
 F3 *see* formant
 F4 *see* formant
 finality 10, 192
 Firthian Prosodic Analysis 130
 flapping 131
 Florentine *see* dialect
 focus 181–182
 formant 27, 74–75, 78–79,
 81–82, 87, 101, 120
 F1 *or* first formant 26, 67, 75
 F2 *or* second formant
 26–27, 67, 75, 78–79, 81, 151
 F3 *or* third formant 74–75,
 78–79, 81, 87
 F4 87
 formant energy 78, 81
 Franco-provençal *see* dialect
 French 3–5, 24, 28, 31–36,
 38–39, 44, 49–51, 172, 174
 frequency 4–5, 20, 31–32,
 34–46, 48–52, 74, 76–77, 79,
 101, 129, 156, 160, 192
 acoustic frequency *or*
 frequency range *or*
 frequency region
 74, 79, 87
 fricative 41, 75, 106, 109–111,
 154–155, 159, 189, 196, 200
 semifricative 159
 sibilant fricative 41, 110
 fronting 19, 26–27, 66–67
 th-/dh-fronting 66–67
 functionalism 3
 future marker 17, 21–22
- G**
 gemination 142–143, 148, 151
 gender 5, 9, 27, 34–35, 62, 65,
 84, 129, 148, 151
 generative 31, 34, 50
 post-generative 34, 50
 Georgian 200
 German 3, 10, 27, 160, 172,
 189–191, 197, 199–201
 gesture 70–72, 74, 81–82,
 84–86, 88, 110–112, 116–117,
 119, 121, 125–126, 128–130
 raising gesture 81–82
 residual gesture 112, 119, 129
 retraction gesture 81
 tongue root gesture 81
 tongue tip gesture 82, 86
 velar gesture 112
 gestural asynchrony 81–82
 glide 20, 110, 123
 glottal 11, 61, 74–75, 86, 103,
 109, 115–118, 121, 127, 131, 190,
 192–193, 195–202
 glottal stop *see* stop
 glottalisation/glottalization
 98, 102, 112, 115–118, 123, 131,
 190, 197, 201
 junctural glottalisation
 197, 201
gorgia toscana 8, 143, 146, 155,
 158, 161, 163, 167
 gradience 118–119, 128, 130, 132
 Greek 25, 90
- H**
h aspiré 32
- I**
 identity 45, 62, 65, 144, 158,
 163, 195
 Scottish identity 62
 social identity 158
 idiolect 3, 18, 140
 imitation 6
 implosive 197
 index 8, 69, 137–142, 144, 146,
 152, 155–156, 158–163
 sociophonetic index 8, 137,
 139–142, 144, 152, 155–156,
 158–159, 161–163
 indexicality 89, 92, 98–99,
 140–141
 social indexicality 98–99
 indicator 116, 119, 126–127, 163
 infinitive 143, 152, 180
 oxytone infinitives
 146, 152–154, 161, 167
 ingressive 197
 Intangible Cultural Heritage
 10, 170, 184
 interaction 5, 7, 10, 18, 69, 90,
 92, 101, 114, 125, 127–130, 132,
 175, 192–193
 interactional 2, 10, 126–127,
 189, 192–193
 interview 24, 67, 99, 127,
 170, 174

- sociolinguistic interviews
99, 127
- ipercorrettismo* 160
- Italian 8, 25, 27–28, 137–139,
141–144, 147–148, 150–152,
154–155, 157, 159–161, 164,
172, 177
- Standard Italian 142–143,
152, 155, 172
- Tuscan Italian 137, 139, 141,
148, 167
- see also* dialect
- L**
- labialisation 107
- lateral 112, 144, 151–152, 156,
158, 167
- length 116, 119, 142–143,
149, 172
- lengthening 82, 142–143,
149, 181
- lengthened 107, 111, 116, 119,
122, 124, 126, 149
- pre-pausal lengthening 181
- vowel lengthening 142
- lenition 61, 102, 105–109,
111–112, 114, 116, 124–126, 131,
137, 143, 160
- lexicalization 153
- Lexical Phonology 7, 98,
109, 130
- liaison 4–5, 12, 31–32, 34–54
- liaison environment 35–36
- linguistic geography 2, 7
- liquid 109, 110
- loudness 180–181
- logarithmic scale 37
- M**
- map-task 184
- marker 17, 21–22, 34, 148, 158,
163–164
- media 6, 66–70, 91–92, 163
- broadcast media 6, 66–67,
70, 92
- Glasgow Media Project 67
- media influence 67–70, 91
- merger 17, 22–23, 61, 74,
150–151
- near-merger 150
- metaphony 177–178
- middle-class *see* class
- mimicry 85
- mnemonic *see* storage
- model 1, 4, 9, 24, 31–33, 39,
62, 67–68, 89, 91, 97–98, 124,
128, 130–132, 137, 139, 141, 144,
154, 161–162, 172–173, 175, 184,
197, 200
- exemplarist *or*
- exemplar model 39
- 'hybrid' model 89
- probabilistic *or*
- stochastic model 32
- regression model 67
- Rothenberg's aerodynamic
model 197, 200
- usage-based model 1, 4, 12,
52, 54
- morphophonology 130
- N**
- nasal 41, 44, 110, 112, 117, 119,
124, 149, 197
- network 158, 167, 175
- non-rhoticity *see* rhoticity
- norm 9, 21, 62–63, 67, 127,
158–159, 175, 202
- covert norm 158
- overt norm 158
- O**
- obstruents 102, 109, 142
- onset 34, 66, 73, 91, 151, 191, 197
- oral archive *see* archive
- P**
- palatalization 144
- parent 4, 17, 19–22, 24–26
- English speaking parents 25
- parental language *or* speech
or system 3, 19, 21, 23
- pause 10, 105, 110, 115, 117, 126,
181–182, 191–192
- floor-holding pauses 10, 192
- perception 1, 6, 52, 73, 85, 89,
92, 124, 128, 130–131, 140–141,
150, 161
- PFC *see* Phonologie du
Français Contemporain
- pharyngeal 67, 75, 79
- pharyngealization/
pharyngealisation
62–63, 75, 81–82
- pharyngealized/
pharyngealised 62, 72,
75, 77–78
- Phonologie du Français
Contemporain (PFC) 4, 12,
31, 33–54
- phonologisation 129
- pidgin 21
- pitch 180–181
- plosive 105, 109, 146, 159–160,
189, 191–193, 195, 197–201
- Polish 25
- postalveolar 61–63, 66, 73
- postvocalic 60–61, 63, 65–68,
72–74, 81–83, 85, 88–89, 91,
101, 143, 146–147
- power-law 4, 31, 35, 37–39
- pressure 3, 189, 196–202
- intraoral pressure 189, 197,
199–201
- supraglottal pressure 197
- prestige 8, 62–63, 140, 142, 156,
158–159, 168
- covert prestige 62–63, 168
- overt prestige 159
- prestigious 62, 157–160
- prosody 131
- prosodic group 50
- Q**
- quality 67, 73–74, 156, 161, 201
- vowel quality 74, 156, 201
- questionnaire 67, 184
- R**
- rafforzamento* *or* *raddoppiamento*
fonosintattico 143, 146, 148
- raising 19, 25–26, 81–82, 85–87,
112, 144, 197, 201
- see also* gesture
- rate *see* speech rate
- 80, 83, 100, 102, 107, 109, 114,
126–127, 130
- reading 34, 61, 63, 68–69, 193
- text reading 34
- word list reading 34, 68,
69, 193
- real time 9–10, 26, 91, 170

- reduction 17, 74–75, 78, 102,
108, 131, 159, 168, 197, 200
amplitude *or* energy
reduction 74–75, 78
cluster reduction 102, 108
release 102–104, 106, 109, 112,
117, 125, 192, 195, 197–201
inaudible release 102–104
plosive release 197–199
silent-released 102
strong release 102
unreleased 102–105, 112
weak release 102
repertoire 8, 18, 50, 139, 142
representation 1, 3, 12, 20, 47,
50, 60, 69, 80, 86–89, 91, 99,
109, 128, 139–141, 147, 176–177,
196, 199
abstract representation
50, 139
acoustic representation 70
articulatory
representation 88
autosegmental
representation 147
cognitive representation 1
mental representation
12, 59–60, 87, 89
phonetic representation
70, 79, 88, 177
phonological
representation 3, 89, 141
psychoacoustic
representation 79
retroflexion 87
rhotacism 144
rhotic 5–6, 60–63, 65–66, 68,
70–71, 73–75, 78–79, 82–87,
89, 91, 99, 152
rhotic-derhotic continuum
see continuum
rhoticity 61–63, 65–67, 70, 75,
79, 82–83, 85–86, 88–89, 91
non-rhoticity 62, 65–67,
70, 91
Romanian 131
rules 12, 32, 40, 97, 99–100,
105, 109, 128–129, 131
categorical rules 99, 127
variable rules 97, 99–100, 135
Russian 151
- S
salience 87, 157, 163–164
sandhi 143, 148
schwa 32, 34, 125
Scots 60, 62, 66, 92
sex 126, 168, 174, 183
sibilant *see* fricative
sign 139
social class *see* class
sonorant 112, 142–143, 149–150
sound archives *see* archive
speaker-hearer chain 88
'speaker-hearer triangle'
87–88, 91
speech rate *see* rate
speech style *see* style
speech archive *see* archive
spelling 34
spike 104, 159
spirantization 143, 148, 155
SSE *see* English
stance 28, 69
stereotype 89–90, 149, 158, 161,
163–164
stigma 8, 158, 164
stop 6–7, 41, 44, 61, 86, 98–107,
109–111, 115–118, 120–121,
124–127, 130–131, 143, 148, 150,
155, 159–160, 191, 195, 197–201
alveolar stop 41, 44, 102, 120
coronal stop 6–7, 98–99,
105, 107, 116
glottal stop 61, 86, 115, 117,
127, 197
word-final stop 6–7, 98,
100–105, 115, 118, 124–127,
130–132, 191
storage 4–5, 12, 32, 40, 50
mnemonic storage 50
stress 21–22, 148–149, 178,
180–181
stress clash 149
stressed 62, 65, 73, 82–83,
107, 114, 142, 148–149, 172,
177–178, 191
unstressed 68, 82–83, 114,
142, 147–148, 152
stricture 196, 200–201
style 5–6, 8, 31, 52, 62, 68, 81,
142, 156, 158, 163, 169–170,
174, 183
- formal style 62, 156, 174
speech style 5–6, 8, 68, 81,
169–170, 183
stylistic analysis 183–184
stylistic shift 63
sub-standard 172
syllable 26, 34, 65, 73–74,
82–83, 142, 144, 149, 151–152,
154, 161, 175–178, 191
syllabification 34
symbolon 90
- T
tap 60–61, 63, 72–74, 76, 83, 85
target undershoot 108, 110, 117
timing 79, 81–83, 85–86, 112
gestural timing 79, 81, 85
tongue 6, 60, 64, 70, 79–87,
112, 125, 197
tongue bunching 85, 87
tongue configuration
79–80, 83, 85–86
tongue movement 64, 80
tongue shape 80–81, 83
see also gesture
transcription 5–6, 33–34, 64,
70–72, 74, 81, 88, 99, 103
auditory transcription
64, 70–72, 74, 81, 88
transillumination 11, 199, 202
trill 60–61, 74, 76
truncation 32, 143, 152, 195
turn-finality 192
turn-holding 193
turn-taking 192
turn-transition points 127
Tuscan *see* dialect
Tyrrhenian-Adriatic
corridor 177
- U
Ultrasound Tongue Imaging
(UTI) 6, 60, 70, 79–82,
84–87
undershoot
see target undershoot
UTI *see* Ultrasound Tongue
Imaging
uvular 24–25, 71–72, 74–75
uvularization 75
uvularized 72, 75, 77

- V
- variable rule *see* rule
- variation 1–12, 17, 28, 31–32, 34–35, 48–50, 52, 59–61, 64, 67, 69–71, 73–74, 81, 86–89, 91–92, 98, 108, 124, 127, 131, 137–142, 144–146, 149, 151, 153–154, 157, 161–162, 164, 172–175, 179, 183–184, 191–192, 201
- diatopic variation 34, 161
- individual variation 8, 17, 179
- intra-speaker variation 183
- intratextual variation 183
- phonological variation 31, 89
- social variation 9, 32, 69
- velarization 8, 144, 146, 151–156, 158, 160
- l*-velarization 8, 144, 146, 151–156, 158, 160
- velum 112, 196
- vocalisation/vocalization 62, 66, 71, 109, 112
- l*-vocalisation 66, 109, 112
- voice quality 95
- creaky voice 112, 115, 119
- see also* quality
- voicing 75, 78, 81–82, 85–86, 102, 118, 122, 124–125, 150, 159, 200
- s*-voicing 159
- vowel 9, 19–20, 23, 25–27, 34, 36, 51, 62–64, 66–68, 71–75, 77–78, 81, 83–84, 99, 101–103, 107, 109–110, 114, 116–117, 124, 127, 142–144, 147–149, 151–152, 156, 160–161, 172, 176–183, 191, 193, 197, 201
- central-low vowel 178
- front mid-high vowel 179
- Italian vowel system 142
- mid vowels 177, 183
- pharygealised/uvularised vowel 62
- plain vowels 62, 68, 72, 74–75
- vowel lengthening 142
- vowel raising 144
- vowel tracks 74
- W
- weakening 79, 81, 83, 91, 105–106, 146–147, 161
- wordlist 62–63, 65, 67, 69, 71, 75
- working class *see* class
- Y
- Yiddish 25–26